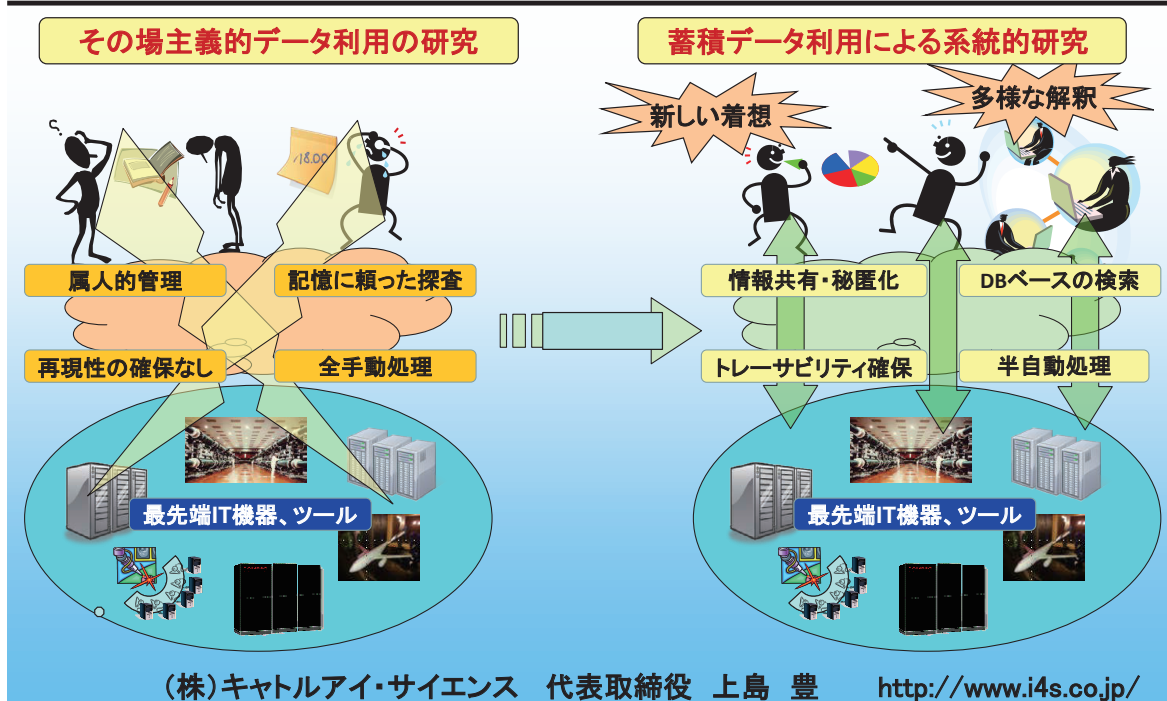


蓄積されたEFD/CFDデータの価値と有効活用に関する提言



目次

- 1) DAHWIN概要
- 2) 世の中の蓄積（ビック）データの状況
- 3) 航空分野における蓄積データの状況
- 4) DAHWINによるEFD/CFD蓄積データ
- 5) 他分野R&D部門における蓄積データの状況
- 6) EFD/CFDの蓄積データで何がかわるか？
- 7) EFD/CFDの蓄積データ活用における課題
- 8) R&D部門での蓄積データ活用の事例
- 9) まとめ

会社概要



研究所で培ったノウハウを結集し、R&D統合プラットフォームを製品化し、幅広くR&Dを支援しています。

社名の由来



・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



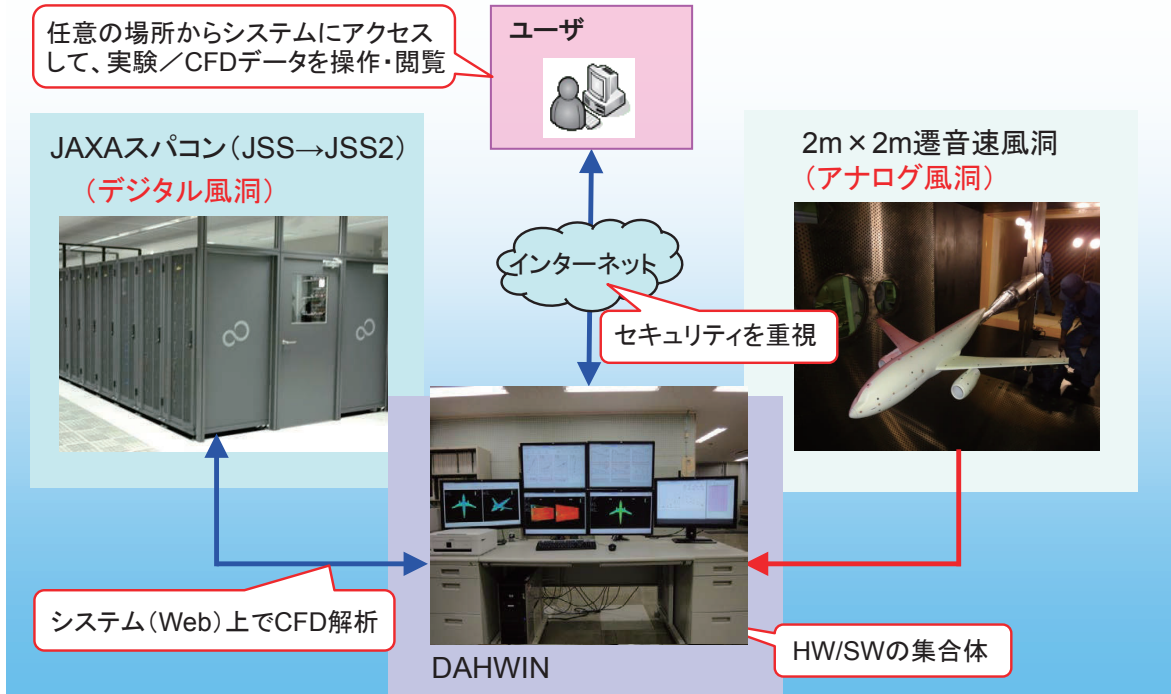
1) DAHWIN概要

DAHWIN概要

引用: 2013.04.17_DAHWIN完成披露会資料



DAHWINは、EFDとCFDを相互補完的に実施し、データを蓄積、データベース化している。

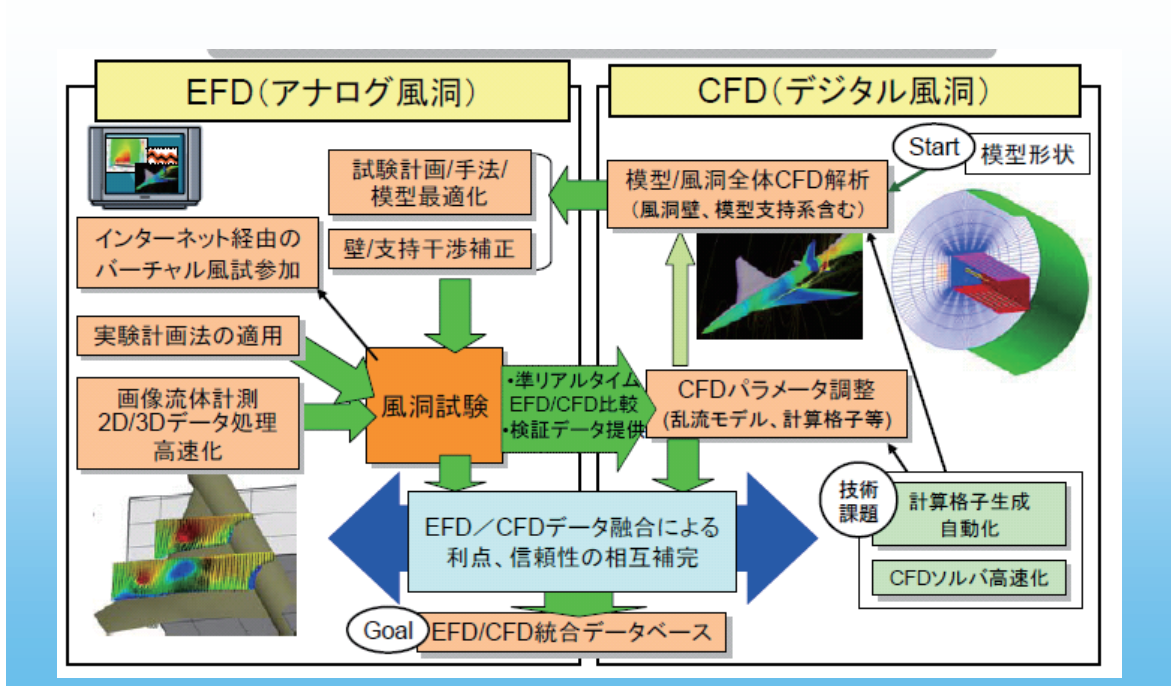


DAHWIN概要

引用: 2013.04.17_DAHWIN完成披露会資料



実験弱点(壁・支持効果等)と計算弱点(モデル限界、精度等)を相互補完！ 実験変更決定も迅速に！



DAHWIN概要

**DAHWINは、様々な機能を実現するだけでなく、
精緻なEFD/CFDデータベースを自動的に構築！**

DAHWINの様々な機能

- 1) 圧力孔配置支援
- 2) 風試条件設定支援
- 3) 模型サイズ決定支援
- 4) 支持装置選択支援
- 5) 風試セッティング支援
- 6) 風洞壁干渉補正
- 7) 支持補正干渉補正
- 8) 計測反映CFD解析
- 9) 風試/CFDデータ不確かさ解析・最尤値推定
- 10) パラメトリックCFD

**=> これらEFD/CFD相互補完機能を使った実験を介し、
EFD/CFDデータを比較できる形でデータベース化**



2) 世の中の蓄積（ビック）データの状況

一般消費者環境のデータ活用



一般消費者向けの環境における蓄積（ビック）データの活用状況を見てみよう。



一般消費者環境のデータ活用



一般消費者対象の環境は、ここ10年の高度化が目覚ましく、蓄積データ活用が進んでいる。

一般消費者向けのITサービス/蓄積データ活用

価格比較 : リアルタイムに様々な店舗での価格・サービス比較

購入履歴 : 過去購入した品物、時期、数量を自動記録、閲覧

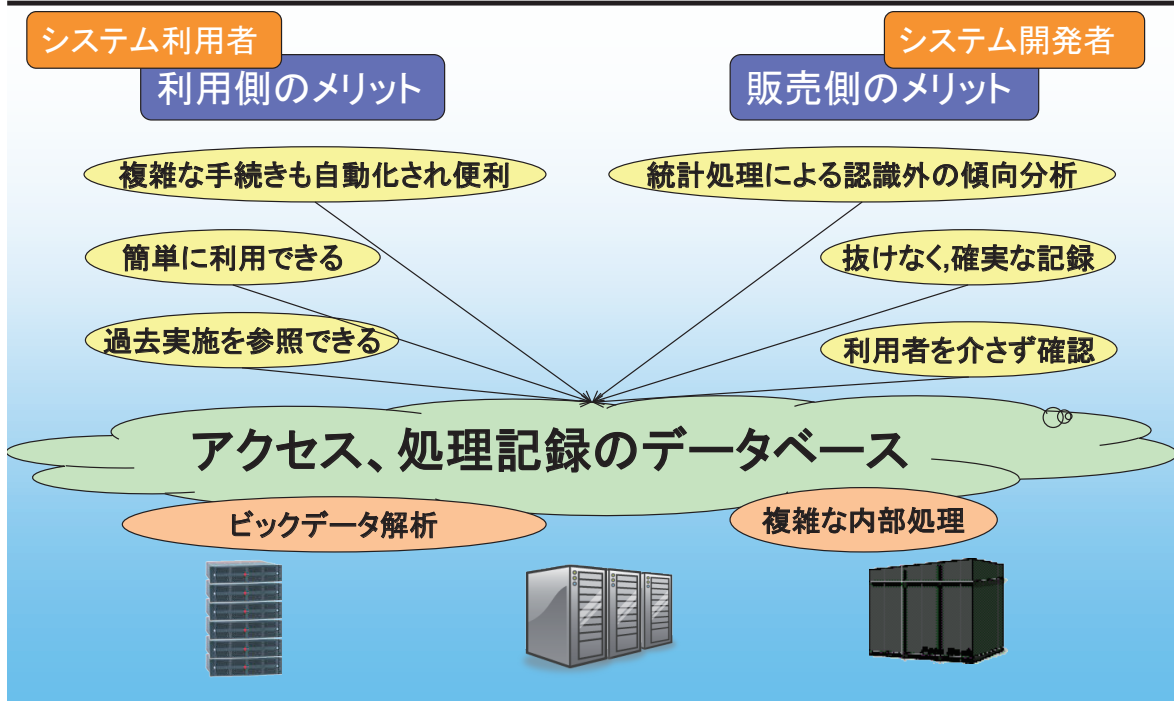
トレーサビリティ: どこでどのように作られたかの情報追跡

質疑応答 : ネットサポート及び利用者同士の質疑応答と公開DB

一般消費者環境のデータ活用



蓄積データ活用は、利用側に対してだけでなく、販売側にもメリットをもたらしている。



R&D部門の環境



現在の一般消費者対象環境とR&D部門の環境を比較してみると！

一般消費者対象環境 V.S. R&D部門の環境

価格比較 ⇔ 結果比較：手動で、属人的（網羅、系統的でない）

購入履歴 ⇔ 解析履歴：手動で、属人的（DB無し）

トレーサビリティ ⇔ 解析再現性保証：手動で、属人的（DB無し）

質疑応答 ⇔ 質疑応答：打合せや個別相談で詳細記録は無し



3) 航空分野における蓄積データの状況



引用: 2013.04.17_DAHWIN完成披露会資料

DAHWINは、EFD/CFDデータの 高い融合と精緻なDB蓄積を実現している。

*現在は運用されていない

	日本	米国	欧州
機関	JAXA	NASA Langley	Univ. Manchester
システム名称	Digital/Analogue Hybrid Wind Tunnel (DAHWIN)	Virtual Diagnostics Interface System (ViDI)	Grid Enabled Wind Tunnel Test System (GEWITTS)*
風試計画設定支援	○	×	×
風試セッティング支援	○	○	×
パラメトリックCFD	○	×	×
風試/CFDリアルタイム比較表示	○	○	×
遠隔地データ確認	○	×	○
風洞/計測機器遠隔操作	×	×	○
データ分析(補正、誤差評価、他)	○	×	×
データ融合(最尤値推定、他)	○	×	×
EFD/CFDデータベース化	○	×?	×?

⇒ ViDI/GEWITTSは、風試/CFD比較表示の利便性向上や、インフラ環境の整備に主眼が置かれており、実験/シミュレーション融合システムとしては不十分

航空分野の状況



DAHWINは、現在の一般消費者対象環境と比べても遜色のないデータ蓄積状態！

一般消費者対象環境 V.S. DAHWIN

価格比較 ⇔ 解析比較：DBベースで網羅、系統的

購入履歴 ⇔ 解析履歴：DBベースで検索可能

トレーサビリティ ⇔ 解析再現性保証：DBベースで保証されている

質疑応答 ⇔ 質疑応答：ある程度はマニュアルで改善余地あり

⇒ ただし、現時点では十分な利活用状態とは言えず、継続的な蓄積データの分析、ユーザ提示法の改善が必要！



4) DAHWINによるEFD/CFD蓄積データ

DAHWINの蓄積データ



DAHWINによるEFD/CFD蓄積データベースは、
入力6(制御4)項目、評価15項目となっている。

入力値

M	気流条件			姿勢角	
	Re [$\times 10^6$]	P0 [kPa]	T0 [K]	α [deg]	β [deg]
0.850	2.27	119.99	322.73	-0.798	0.0
0.850	2.25	119.96	325.38	0.000	0.0
0.850	2.25	119.97	325.57	5.000	0.0

+ モデルデータ

評価値

空力係数														
機体軸							安定軸							
CA	CY	CN	CI	Cm	Cn	CAF	CD	CY	CL	CI	Cm	Cn	CDF	CLF
0.01825	-0.0	-0.00416	0.0	0.15963	-0.0	0.01825	0.01831	-0.0	-0.00390	0.0	0.15963	-0.0	0.01831	-0.00390
0.01837	-0.0	0.06847	0.0	0.13730	-0.0	0.01837	0.01807	-0.0	0.06855	0.0	0.13730	-0.0	0.01807	0.06855
-0.00121	0.0	0.67383	0.0	0.02906	-0.0	-0.00121	0.05749	0.0	0.67138	0.0	0.02906	-0.0	0.05749	0.67138

DAHWINの蓄積データ



CFD蓄積データに関しては、数値計算上の条件
および支持、壁の有無もDB化されている。

CFDデータの追加情報

CFD		収束結果 判定			解析情報								
種別	区分	ΔCL [$\times 10^{-4}$]	ΔCD [$\times 10^{-4}$]	Res [$\times 10^{-6}$]	アプリケーション情報		解析条件		格子情報				
		収束判定値: 1.0 $\times 10^{-4}$	収束判定値: 1.0 $\times 10^{-4}$		格子	ソルバー	粘性	計算回数	半/全	位置	空間セル数	表面セル数	代表長
事前CFD	支持あり	0.020000	0.002000	0.000565	HexaGrid	FaSTAR	SA	32857	半	右舷	12335656	248796	151.310
事前CFD	支持あり	1.800000	0.212000	0.002580	HexaGrid	FaSTAR	SA	260840	半	左舷	33826956	701728	151.310
事前CFD	支持あり	0.030000	0.004000	0.001583	HexaGrid	FaSTAR	SA	220569	半	左舷	34421014	670970	151.310

+

収束済解析結果生データ

DAHWINの蓄積データ



EFDは約13万計測、CFDは1700計算のデータが蓄積され、CFD/EFD~1/30程度である。

EFD/CFDデータの蓄積量

	モデル数	計測数	1モデルあたりの計測数	M,α,βあたりの計測数	M,αあたりの計測数
EFD	61	約130,000	約2,100	約13	約46
CFD	26	約 1,700	約 65	約 4	約 8
CFD/EFD	0.43	↑ 0.013	0.03	---	---

CFDは、EFDの1/30程度実施

CFDは、約3500時間=146日もの連続計算相当



5) 他分野R&D部門における蓄積データの状況

他分野の蓄積データ状況



DAHWINは、他分野のデータ蓄積が多い研究所に肩を並べ、解析も同時蓄積している稀有な存在！

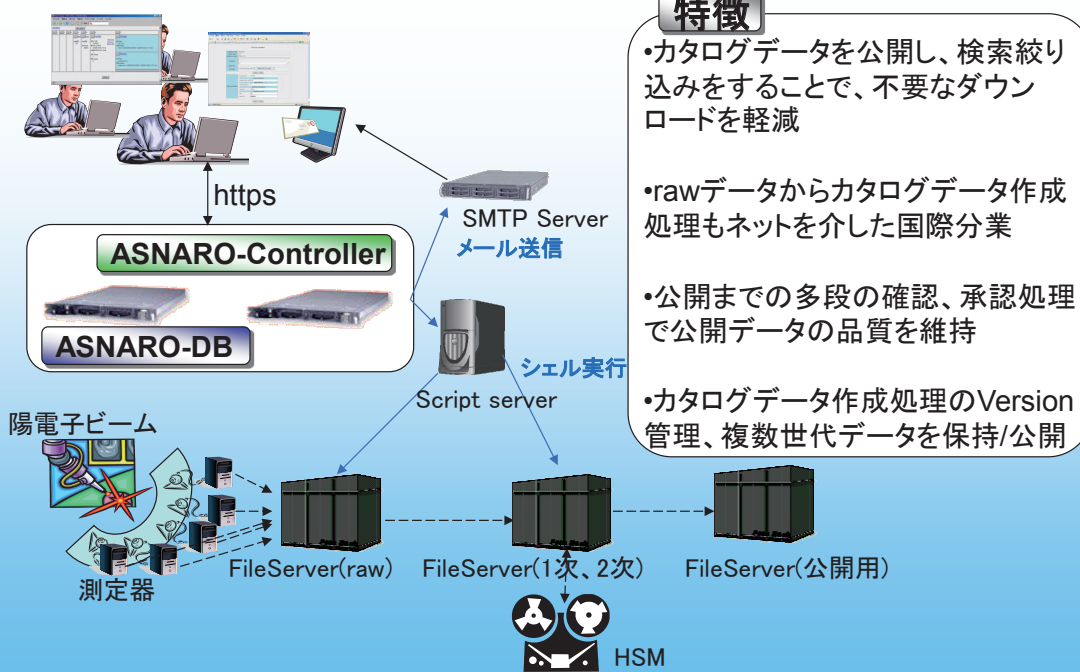
様々な分野のデータベース化されたデータの蓄積量

	蓄積年数	計測数
KEK/素粒子実験	10数年	10万弱
KEK/素粒子解析	10数年	10万弱
KEK/蛋白質構造解析	4年	約10万
すばる望遠鏡/天体観測	10数年	数10万
JAEA-APR/レーザー実験	1年	約500
JAEA-APR/レーザー解析	1年	約3,000
DAHWIN/実験	3~4年	約130,000
DAHWIN/解析	3~4年	約 1,700

KEK/素粒子



数ペタByte、数百万ファイルに及ぶ実験および理論モデル解析結果をデータベース化し、国際的に公開！

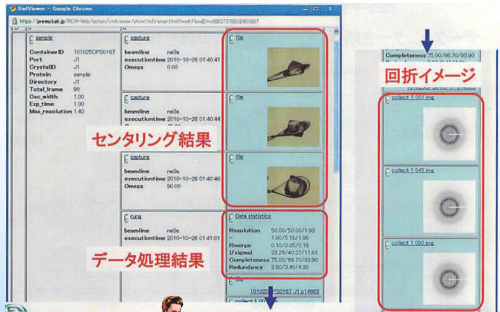


KEK/蛋白質



蛋白質の結晶化及びX線構造解析実験を 堅牢なアクセス制御機構下で、丸ごとDB化！

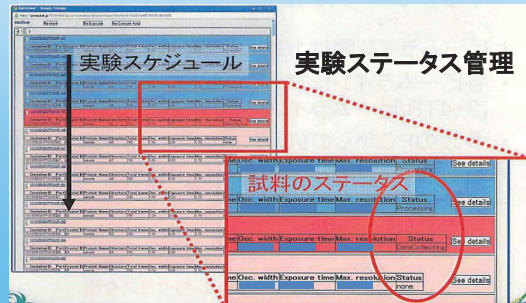
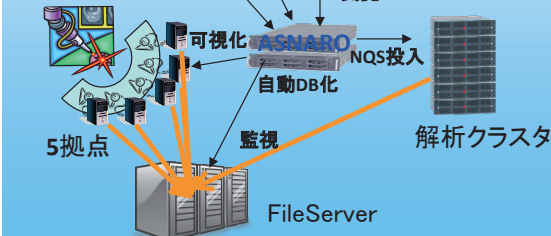
実験結果表示



特徴

- 連続計測が、リアルタイムで解析・可視化
- 遠隔地からも実験に参加可能
- 過去の類似実験との比較が容易
- 新規解析手法の導入が容易
- 蓄積データを用いた特性分析が可能

X線観測施設



すばる/天体



100セット/日の観測データを年々改善される10数段に 及び画像解析処理とともにDB化され、3年後に公開！

RCM System Software ユーザ事例

大学共同利用機関法人

自然科学研究機構

国立天文台

選定理由

■自分たちで進化させられる

システムの外注はコスト的にもシステムの発展的にも難しいので、我々の手で自ら構築する方向で動き出し、当初は、PythonやPerlというプログラミング言語と無料のデータベースアプリケーションを用いて自作する予定だった。
あるとき高エネルギー加速器研究機構の先生方と会議をしている際に、彼らのBelle実験システムで大規模データ管理用にRCM System Softwareというミドルウェアを採用しているというお話を伺った。彼らの実験も大規模で、且つ天文学におけるデータ解析と類似するデータ処理プロセスを含むことが分かり、もしやと思いついてそのソフトウェアの販売元に問い合わせたところ、日本原子力研究機構の元研究者が我々と同様の問題を克服し、そのノウハウを製品化した研究者のためのミドルウェアであることがわかった。

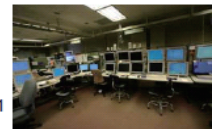
このソフトウェアを我々が採用した理由は、システム構成変更しに追従可能で、且つ天文学者達自らがシステムを持続的に進化させることが可能な柔軟なデータベースとフレームワークをもったソフトであったことだ。

国立天文台 古澤 久徳博士

USER PROFILE

大学共同利用機関法人
自然科学研究機構
国立天文台
〒181-8588
東京都三鷹市大沢2-21-1

<http://www.nao.ac.jp>





6) EFD/CFDの蓄積データで何が変わるか?

蓄積データ活用



**折角作った多数の機能は、蓄積データを使って、
評価、改善（利用停止）をしていく必要がある！**

蓄積されたEFD/CFDデータを使った機能の評価と改善指南

- 1) 圧力孔配置支援
- 2) 風試条件設定支援
- 3) 模型サイズ決定支援
- 4) 支持装置選択支援
- 5) 風試セッティング支援

=> 事前支援機能を十分使っていればどの程度の追加労力でどの程度の改善があったかを過去のデータを使って、定量的に評価する。

=> 追加労力対改善が十分でない場合は、支援機能を強化する。

- 6) 風洞壁干渉補正
- 7) 支持干渉補正

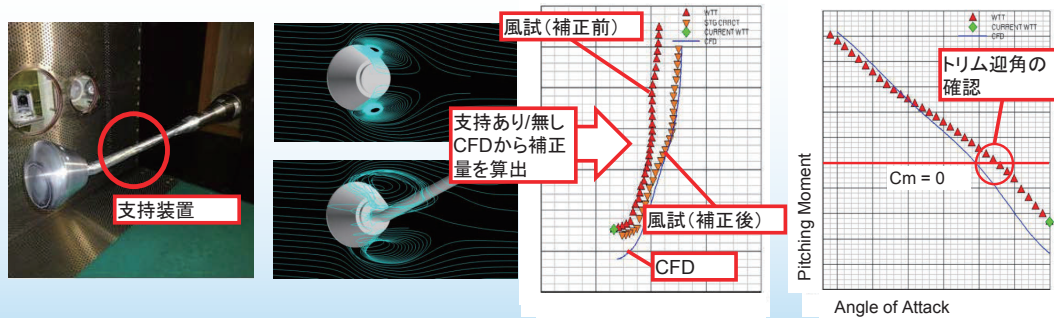
=> 補正前後の差が必要精度より大きいか、そもそもCFDが必要精度内かを明確にする。

蓄積データ活用

引用:2013.04.17_DAHWIN完成披露会資料



回収機能付加型HTV（HTV-R）試験では、 支持装置補正支援が研究効率を大きく改善させた。



- 事前CFD: 全174ケース実施(全所要時間2週間程度)
- HTV-Rプロジェクトチーム事業所(筑波)において、風試状況をモニタリング
 - ⇒ 風試関係者の出張頻度低減、データ確認効率・高速化に貢献
 - ⇒ モニタリング画面によるトリム迎角の確認がシステムにより実現
- 模型支持装置の影響を、CFDにより補正
 - ⇒ 風試中における補正*の実現により、風試データのリアルタイム妥当性確認に貢献

*従来は風試後のポスト処理

DAHWINの特記すべき効果: ・出張頻度の軽減、・風試とCFDの抵抗データの差異の原因が支持干渉であることが即座に確認でき、無駄な追加作業なく試験を続行、・圧力分布切出し機能により風試/CFD両者の圧力分布の差異の原因を風試中に究明(風試後作業削減かつ手戻りなし)

蓄積データ活用



要求精度に比べて、補正機能が有効かを確かめる！ (要求精度は、事前に決めておくべき！)

8)計測反映CFD解析機能

- => 変形前後の差が必要としている精度より大きい(重要)かを明確にする。
- => 変形前後の差が重要であれば、その予測値を事前計算する手続きを確立し、モデル作成時に反映させる。

9)風試/CFDデータ不確かさ解析・最尤値推定機能

- => 不確かさ(エラーバー)が必要としている精度より大きい(重要)かを確かめる。
- => 最尤値に関しては、単純にCFDとEFDの誤差から算出していいのか?を再検討する必要がある。

10)パラメトリックCFD機能

- => 実験後、不足している(実験で急激な変化のある)部分を追加CFD実行をする。

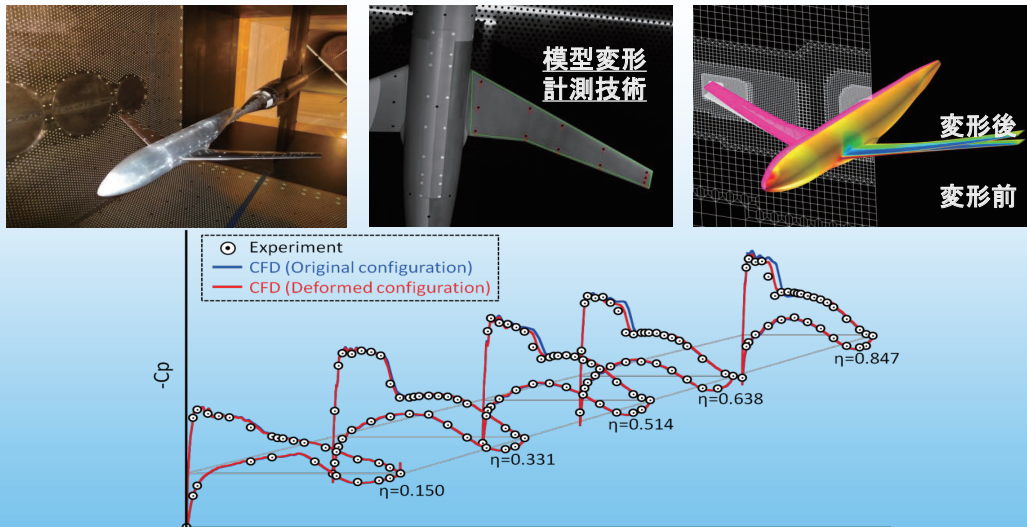
蓄積データ活用

引用:2013.04.17_DAHWIN完成披露会資料



DLR-F6標準模型試験では、計測反映CFD解析により、EFD/CFDの厳密な比較検討が可能に。

■ 模型変形計測反映CFD解析結果



⇒ 風洞試験条件を忠実に模擬したCFDにより、EFD/CFDの厳密な比較検討が可能に

蓄積データ活用



蓄積されたEFD/CFDデータは、実験装置、解析コード、理論モデルの高度化の源泉！

蓄積されたEFD/CFDデータのさらなる活用

- 11) プレ風洞試験における支持装置へのモデル固定の正しさ評価
- 12) 計測器、風洞システムの健全性評価
=> 失敗/成功実験のデータを分析し、推定方法を確立
- 13) 解析結果と実験結果の差の定式化
- 14) 解析コードの高度化
=> 要求精度観点で実験との乖離が大きい所からの改善着手
- 15) 空力係数の近似モデル構築
- 16) 試験飛行体データとの比較による壁・支持装置干渉効果の検証



7) EFD/CFDの蓄積データ活用における課題

データ活用の課題



蓄積データを効果的に活用するには、コミュニティ、組織の理解/評価と融合領域的人材が必須！

蓄積されたEFD/CFDデータを活用することで、多くの重要な新しい知見が得られる。また、データは**一定期間後、公開することが望ましい。**

- 知見を得るのに無限の時間を使っていいとは言わないが、ある程度の時間をかけることを容認、評価する必要がある。

=>追加労力対改善が十分でない場合は・・・
=>変形前後の差が必要としている精度より大きい・・・
=>不確かさが必要としている精度より大きいか・・・
=>失敗/成功実験のデータを分析し、・・・
=>要求精度観点で実験との乖離が大きい所から・・・

- また、データ分析する人材は、多量のデータを扱うセンスと航空分野研究者と意見交換できる融合領域的能力が必要。



8) R&D部門での蓄積データ活用の事例 (ずいぶん昔ですが、自分自身の事例です)

蓄積データ事例



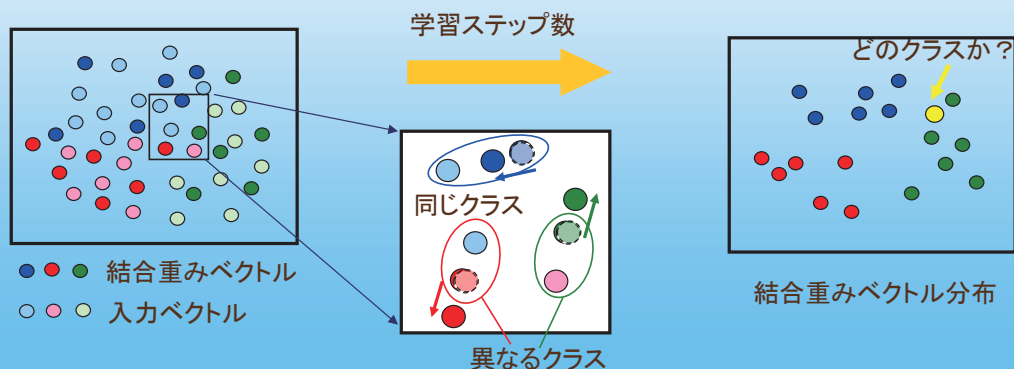
LVQという知識自動獲得技術で、データの分類や最適パラメータ値を探し出す。

学習ベクトル量子化法 (Learning Vector Quantization : LVQ)

分類したいデータの複数の特性指標をひとつのベクトルデータと考える。

様々なデータはその多次元ベクトル空間に分布していることになる。

すでに分類されているデータセットを用いて分類を正しくできるように多次元ベクトル空間を結合重みVectorを使い、分割 (Quantization) を学習 (Learning) する。



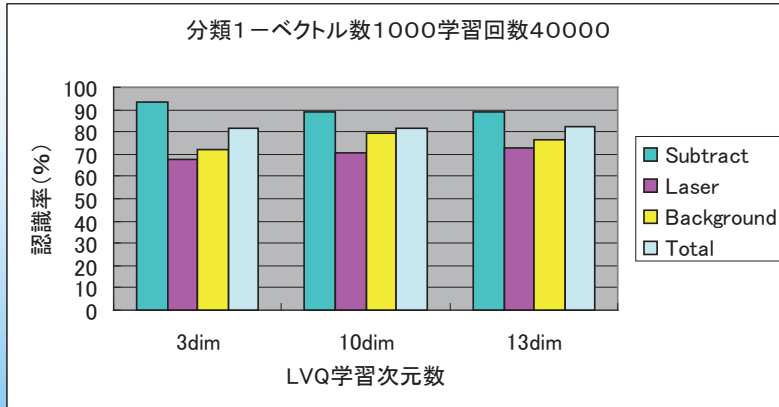
蓄積データ事例

引用:2005.09.21 IPSJ SIG Technical Report MPS 56



LVQをCCD画像情報の統計情報に適用し、90%程度の確率でSubtractデータを分類できた。

サンプル数:Laser:144 Background:86 Subtract:216 total:446



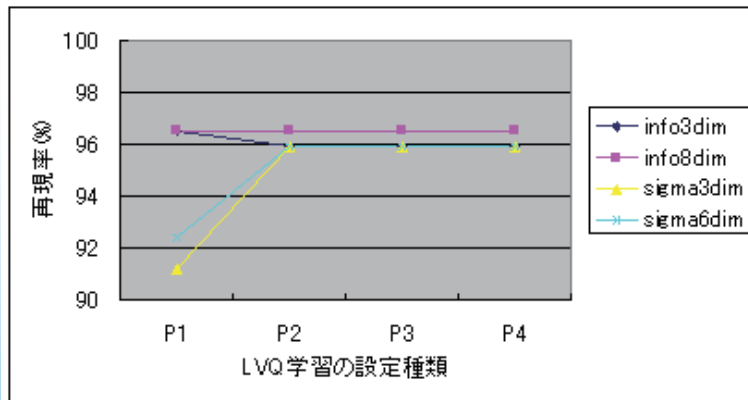
X軸:LVQ学習における次元数(画像統計量) Y軸:LVQ学習後の精度(エラーの小ささ) ピーク値、ピーク位置、平均値、分散など

蓄積データ事例

引用:2005.09.21 IPSJ SIG Technical Report MPS 56



間違ったデータペアを96.5%の再現率(数え落しの小ささ)で判定できた。

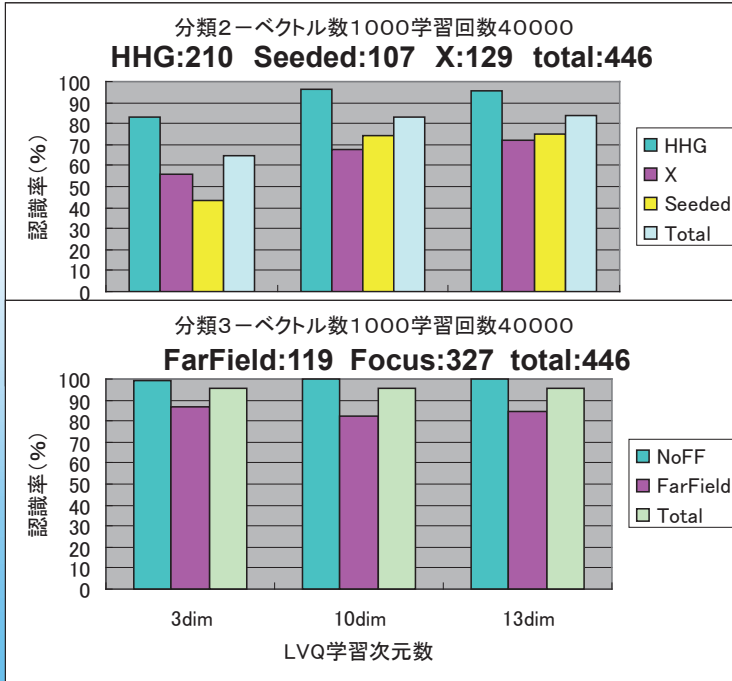


研究者が間違っただータペアを登録しようとした場合、警告する。

蓄積データ事例

引用:2005.09.21 IPSJ SIG Technical Report MPS 56

LVQでの自動分類は、実験の物理的な特性にまで有効である可能性が見出された。



HHG実験のデータの分類は、95%以上の確率で分離

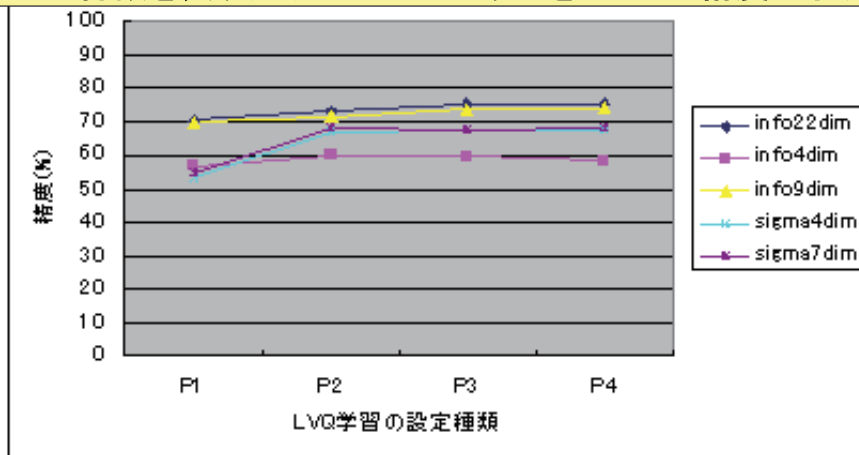
Focus-FarFiledの分類は、非常に高確率で分離

蓄積データ事例

引用:2005.09.21 IPSJ SIG Technical Report MPS 56

2次元イメージから特徴的なXY断面を生成する場合にも、LVQによる学習法が有効である

あるXY断面(断面位置、積分区間)を生成する可視化パラメータが、そのイメージの特徴を表す適切なものかどうかを75%の精度で予測できる。



乱数か何かで可視化パラメータをたくさん発生し(可視化は不要)、2つの可視化パラメータ候補を見つけ出せば、ほぼ確実に研究者が得たい断面像を自動生成できることになる。



9) まとめ

まとめ



EFD/CFD蓄積データは、宝の山だが、有用な情報を引き出すためには、組織の理解と融合領域的人材が必須！

- DAHWINは、他分野のデータ蓄積が多い研究所に肩を並べるデータ蓄積力があり、さらに解析も同時蓄積している稀有な存在！
- 知見を得るのに無限の時間を使っていいとは言わないが、ある程度の時間をかけることをコミュニティ、組織が容認、評価する仕組み、制度が必要。
- また、データ分析する人材は、多量のデータを扱うセンスと航空分野研究者と意見交換できる融合領域的能力が必要。



早急な成果を求められる現代科学研究だが、地道な（個人的志向でなく）積み重ねなくして、明るい未来はない！

キャトルアイ・サイエンスは、R&Dプロセスを革新し、技術者の創造性を支援します！

感性と知性を
情報技術で
革新します
4つの | を科学する



ASNARO

