Experiences in the corporate-wide deployment of modeling technology

Salvador García Muñoz

Paul Schmitz

Pfizer Worldwide R & D

PMET
Process Modeling & Engineering Technology

Pfizer

The Lifecycle of a model in industry

- Models are:
  - Conceptualized
  - Developed and refined
  - Tested and Verified
  - Applied to a problem… over and over
- Implying
  - SME time is invested in running a model over and over and not necessarily focusing on the next big thing
  - Intensive exercise to “lock” model to prevent misuse
  - Training needs for users are very high to handle native software
  - Labor intense bookkeeping of Excel/VBA interfaces
  - Uneven version usage across a corporation
  - Auditability left to the user
  - Suboptimal use of software licenses

Lifecycle of a model in industry

- Specially difficult if…
  - Parameter estimation results need to be handled appropriately
  - Complex computational infrastructure is involved (Linux Cluster)
  - Model versions change frequently
  - Large usage
  - Multi-step process that need to remain in the native platforms:
    Excel → Matlab → gPROMS → GastroPlus → etc
  - High turnaround and transfer of information is only written (simulations done 2 years ago and person is gone?)

The Reality

- Simulation tools are not standard…

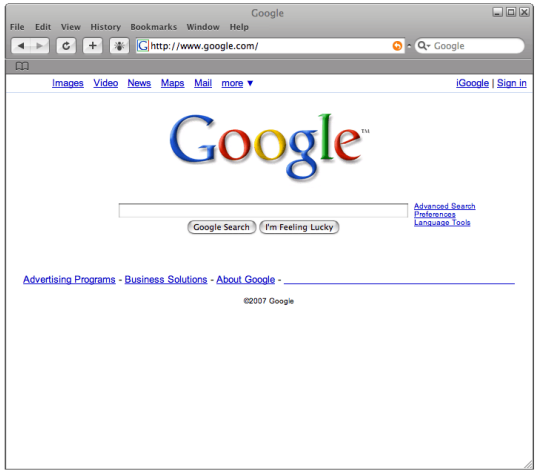
Model	Platform
Solubility Calculations	Cosmotherm Material Studio SAFT….
Kinetic Studies, Reactor Design	Dynochem, Aspen Plus, gPROMS…
Crystallization	gCRYSTAL, Dynochem, Aspen+…
Solid Unit Op's	gSOLIDS, Solidsim, Custom Code, gPROMS, ACM, MATLAB…
Content Uniformity	SimCU, Intellipharm…
Oral Absorption	gCOAS, GastroPlus, SimCyp…

Very likely a numerous amount of in-house built spreadsheets
And the odd code in C++ left by an intern 2 years ago….

C++

The Reality

- Web access is standard…

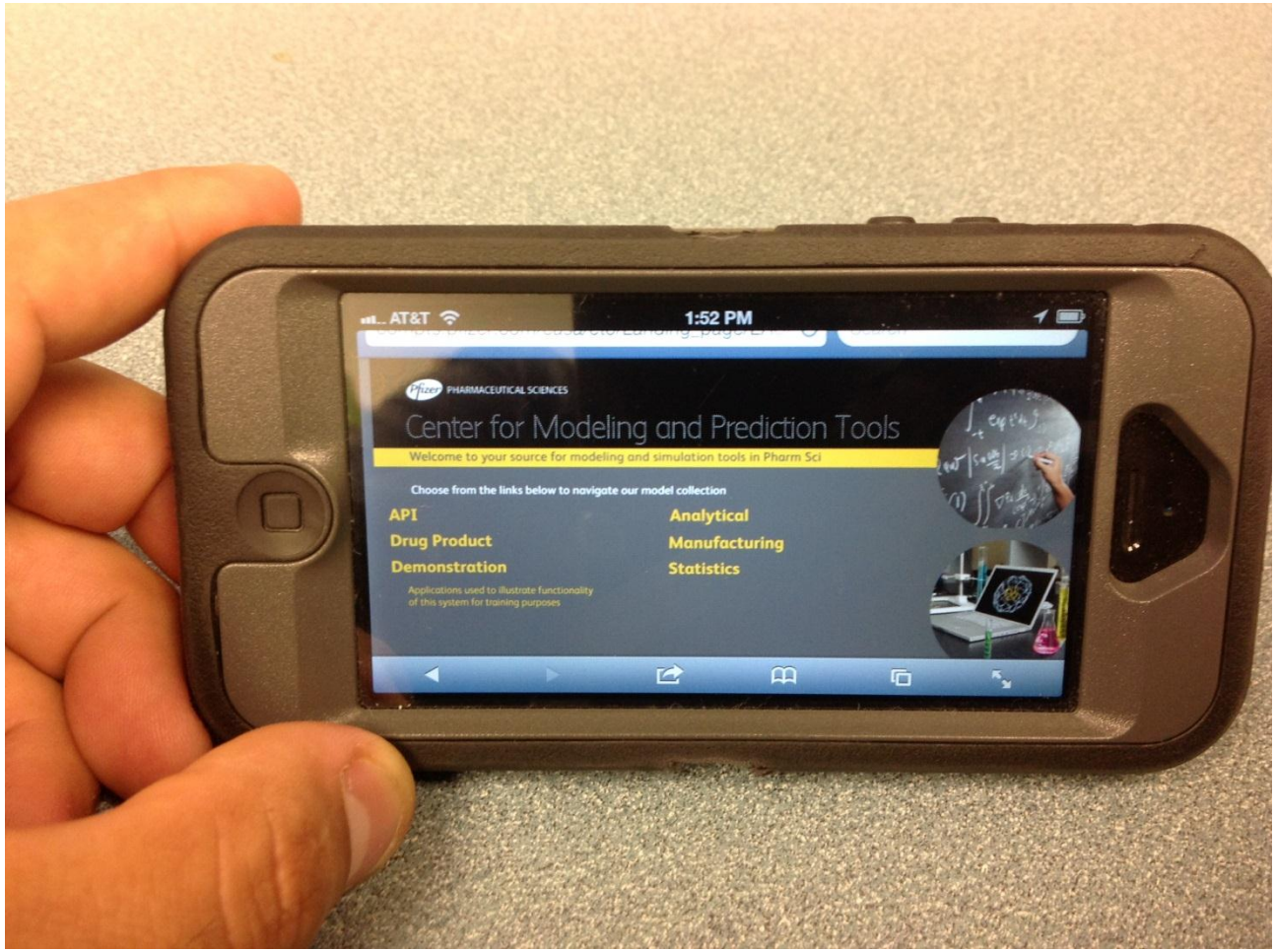


The desired state

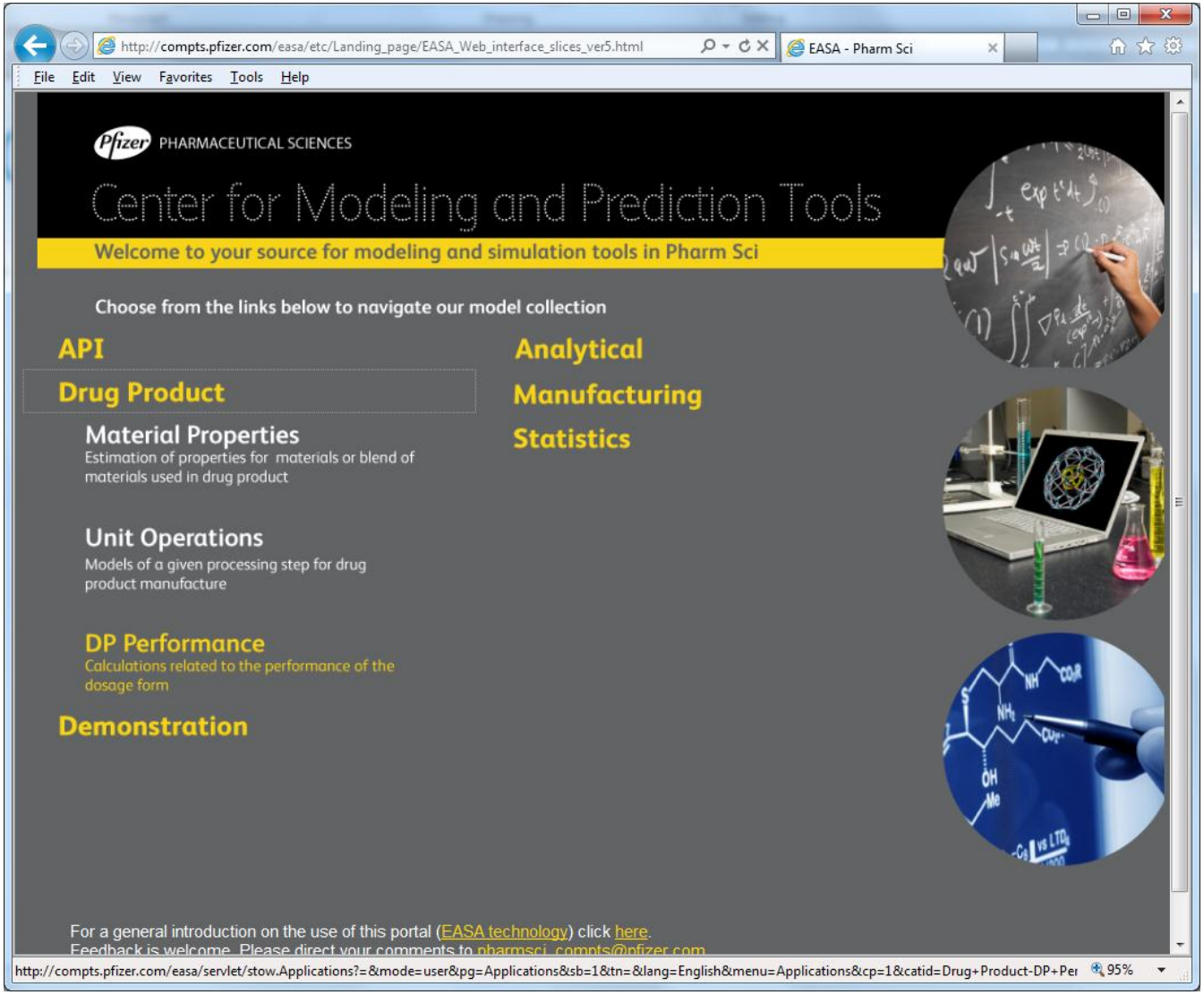
- Centralized one-stop-shop access points for all model **usage** (not development)
- Same “look and feel” to minimize “*where do I click?*” training efforts.
- Auditable space to retrieve inputs/outputs, model version, and all other associate files related to a simulation
- Version Control
- Documentation Central
- While transparently running each model in **its native platform** (even in Linux cluster) with minimum programming effort

Our Solution: Web Accessible

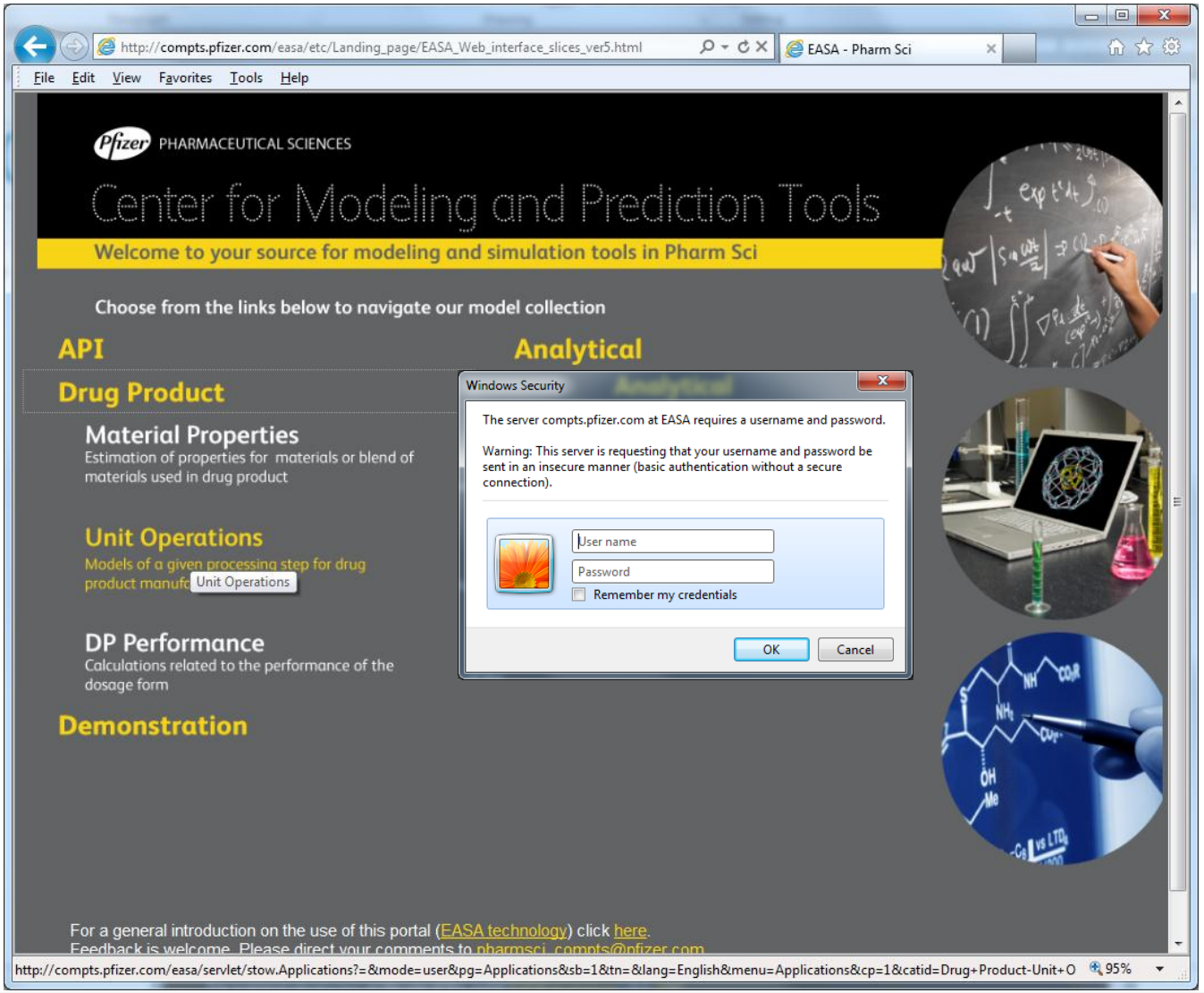
Our Solution: Web Accessible



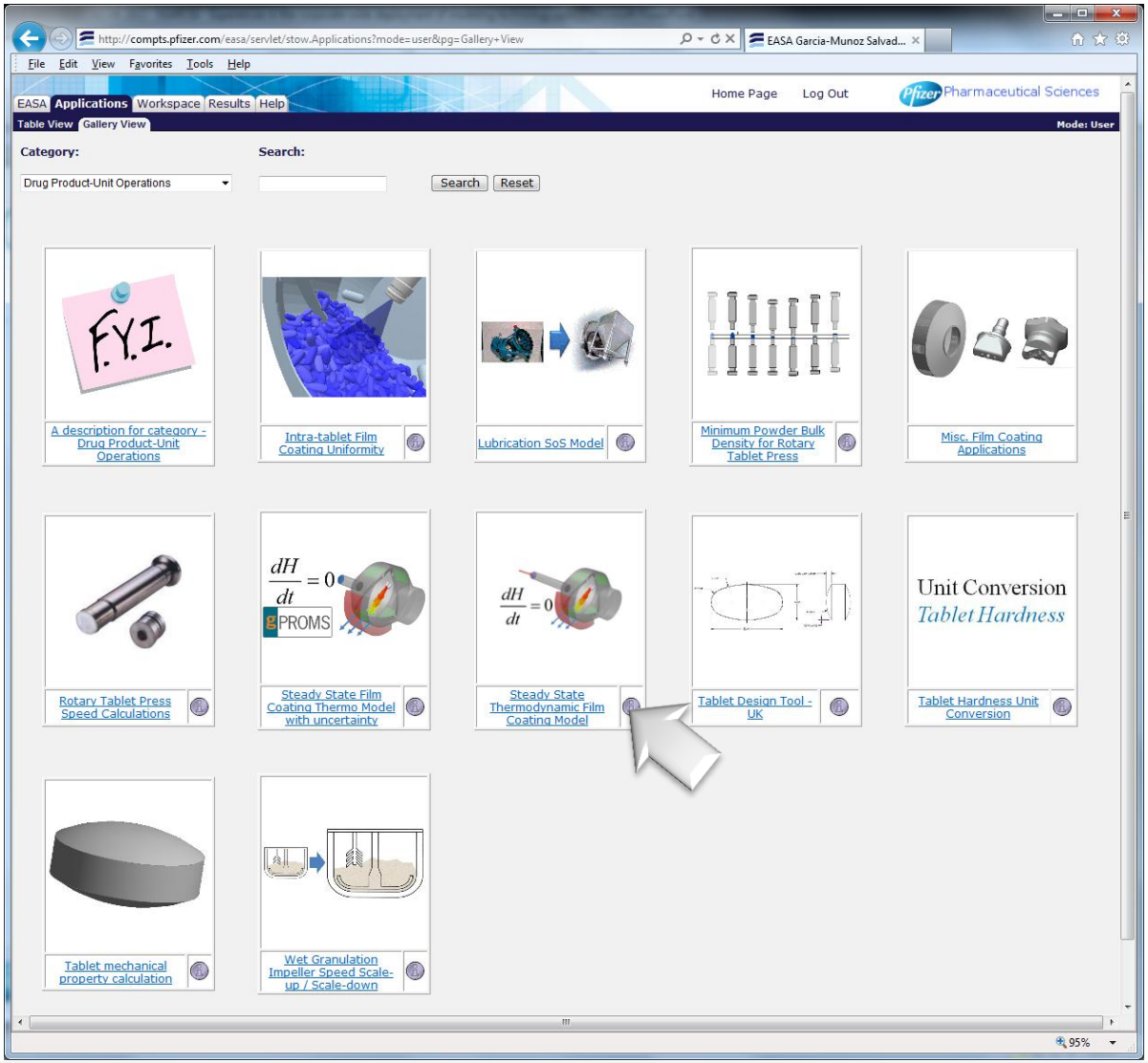
Our Solution: Easy to navigate



Our Solution: Secure

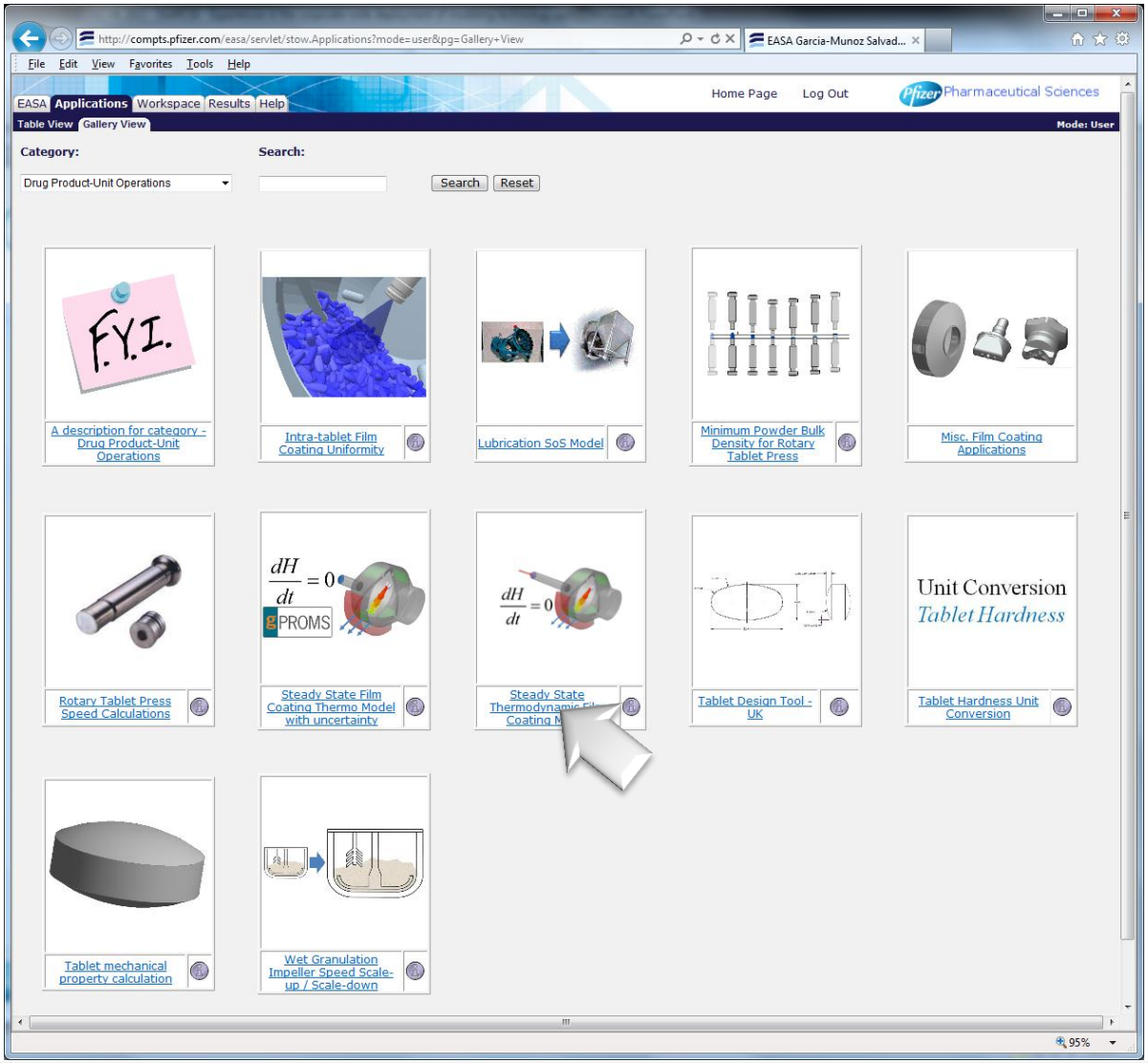


Our Solution: User Friendly

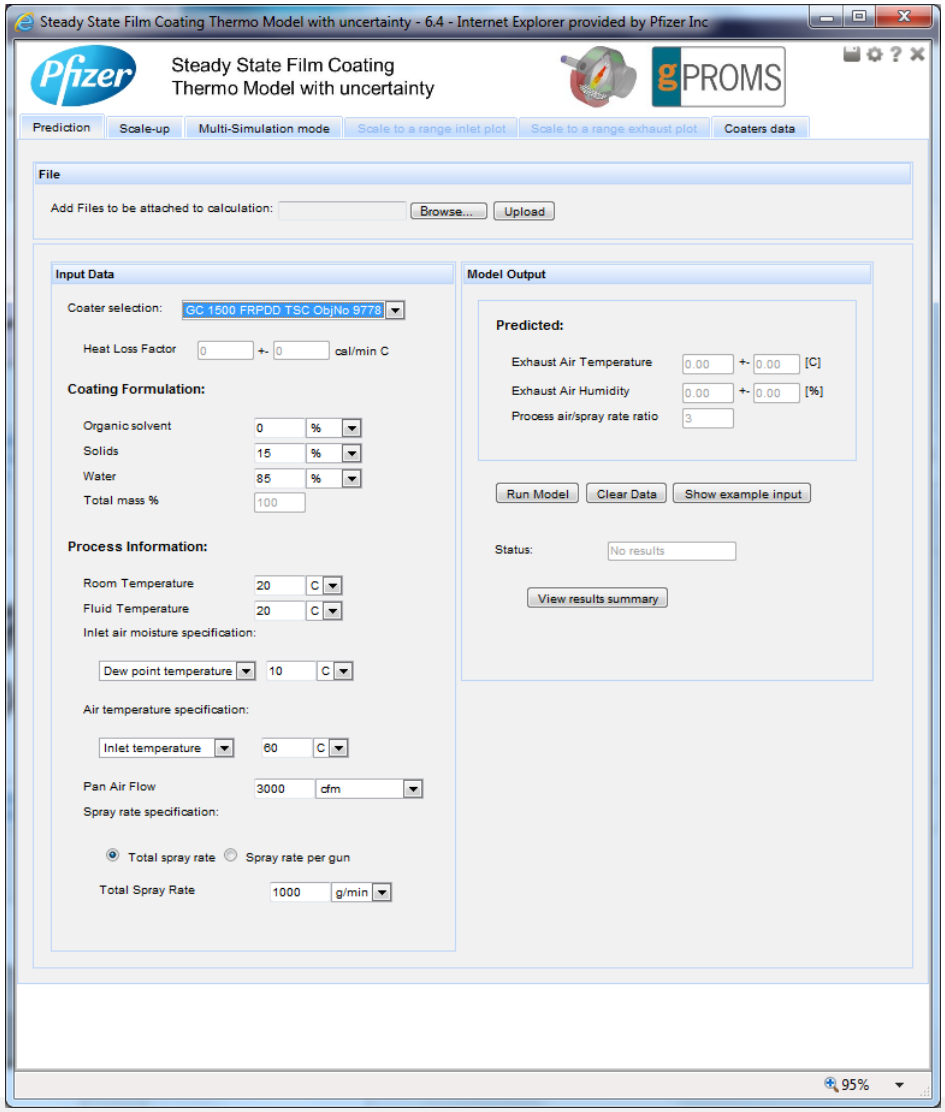


Our Solution: Documentation Central

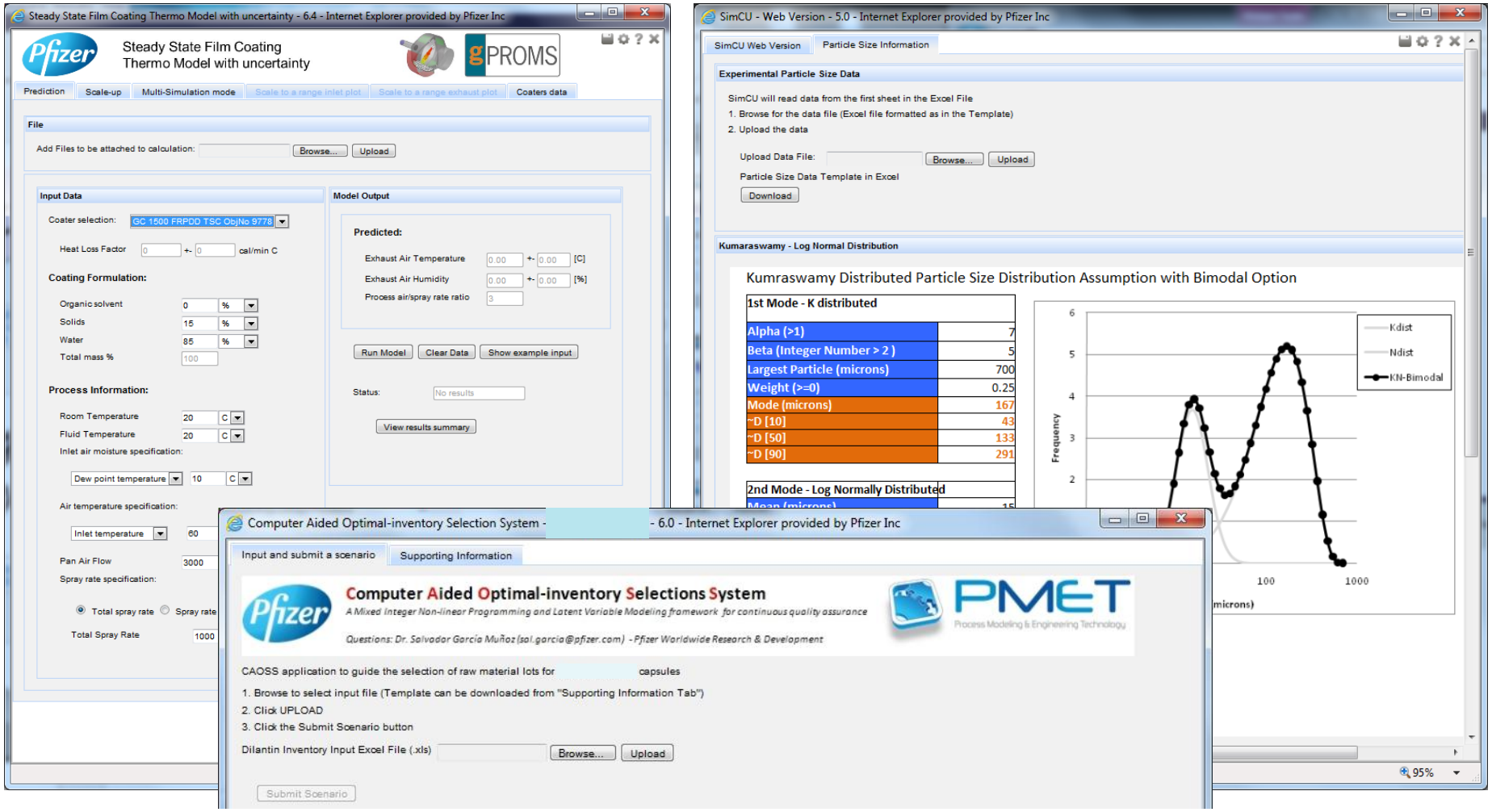
Our Solution: User Friendly



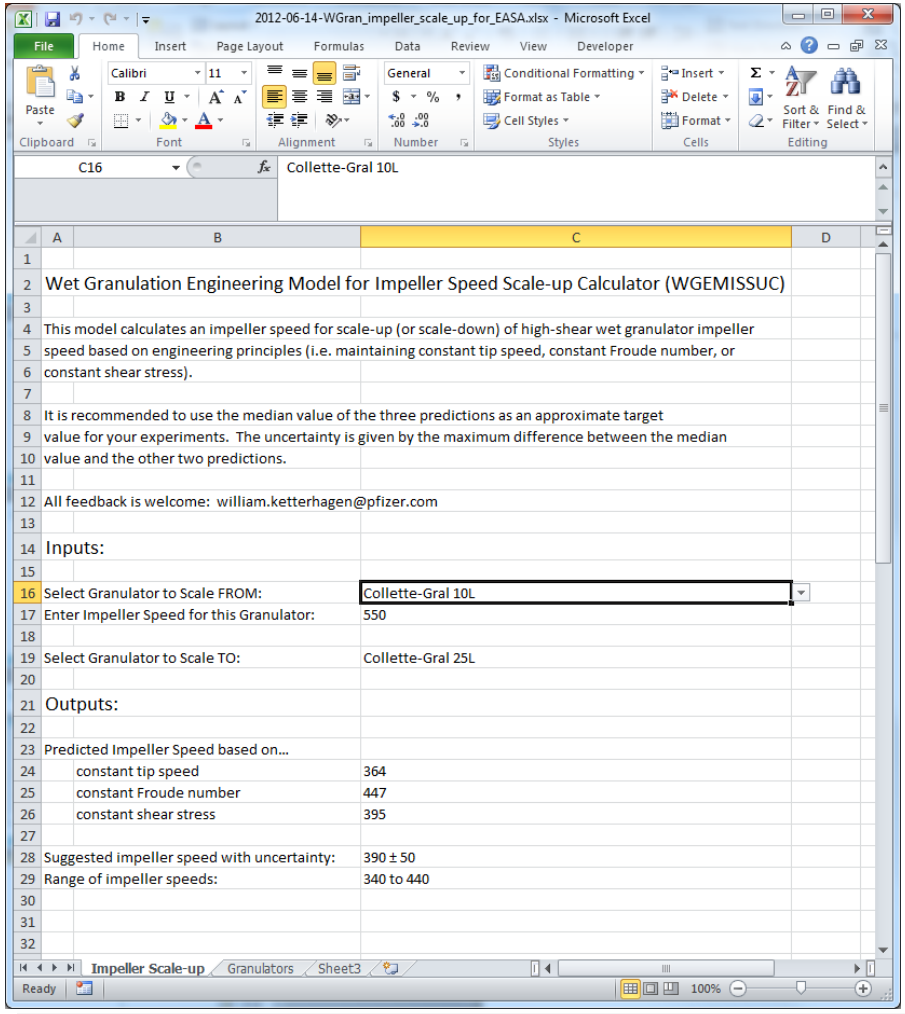
Our Solution: Tailor fit

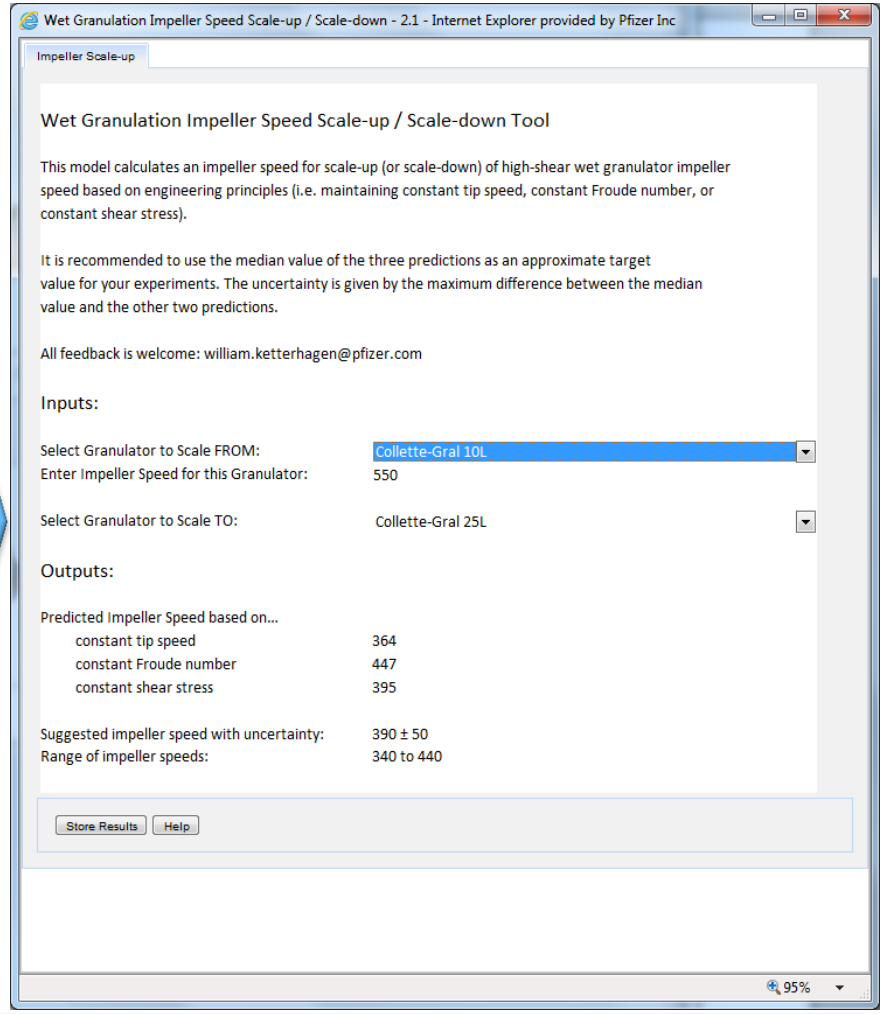


Our Solution: Same “Look n feel”

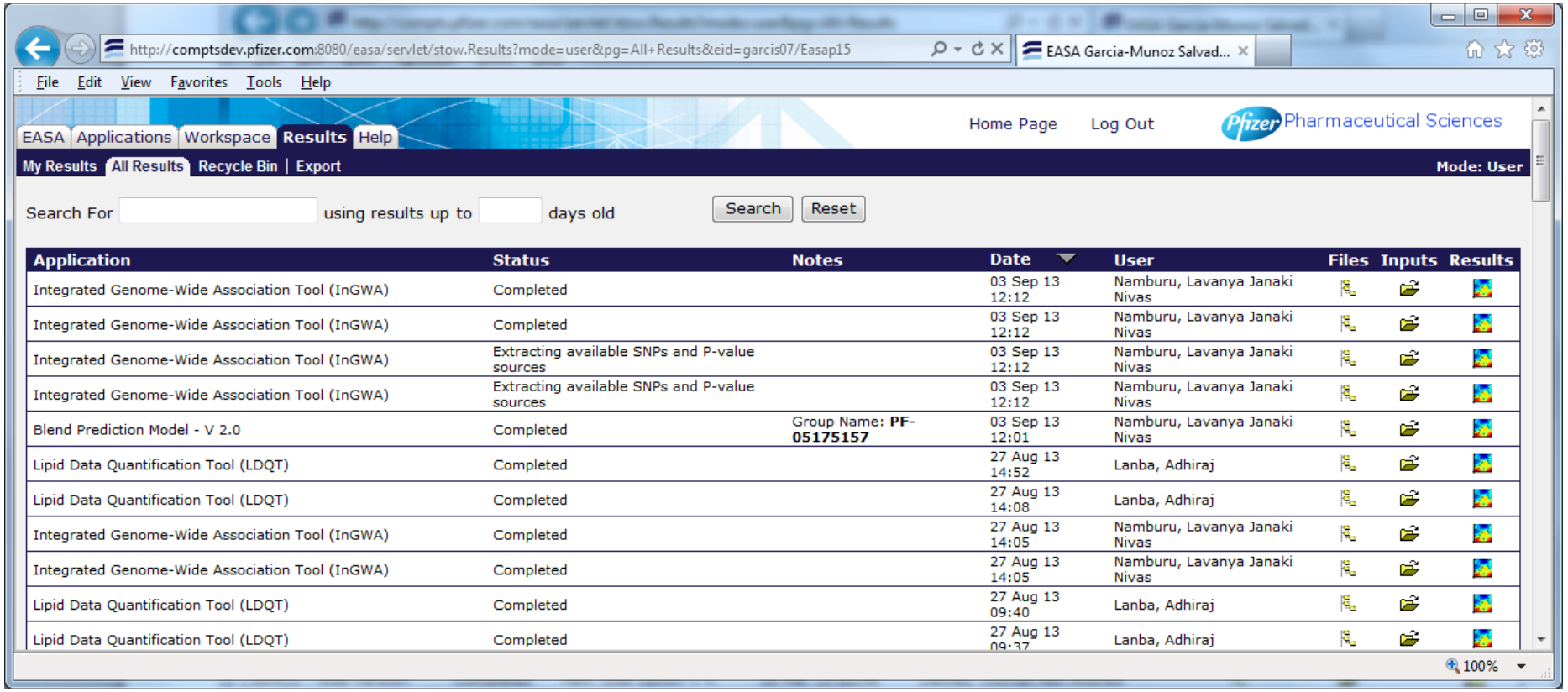


Our Solution: Code-less creation of GUI



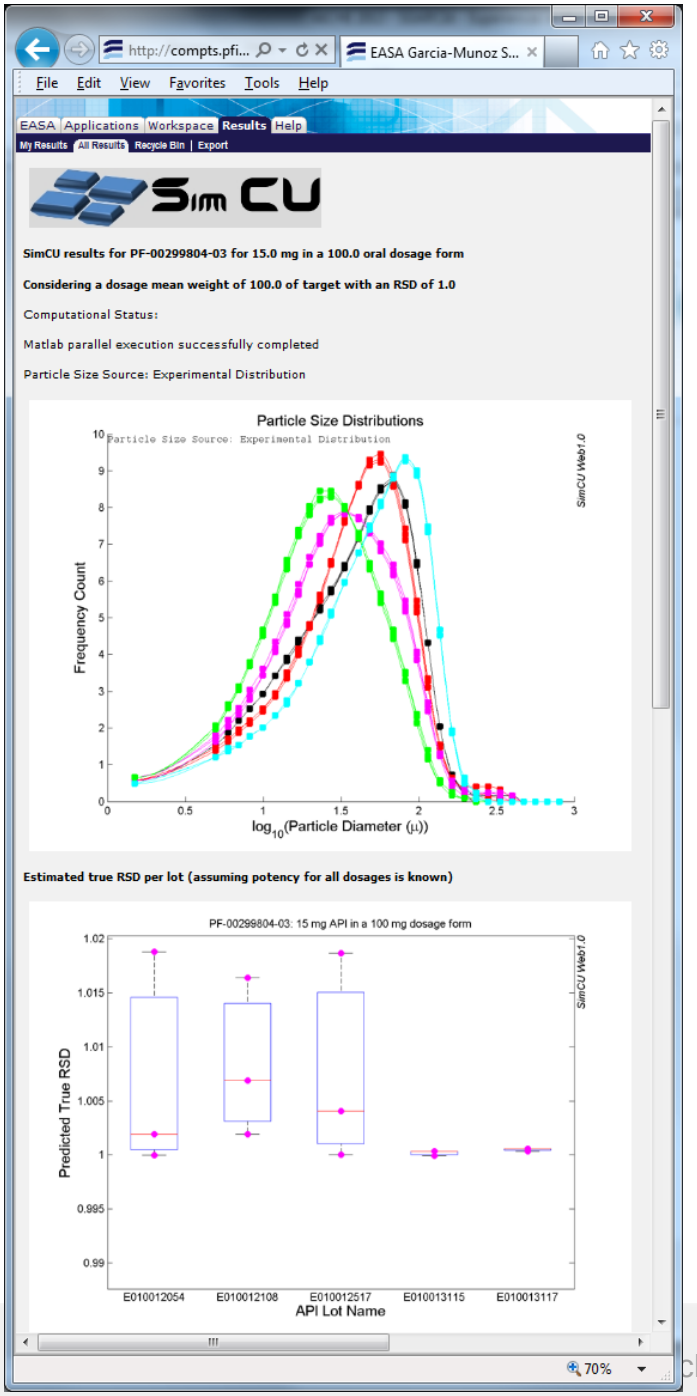


Our Solution: Auditable

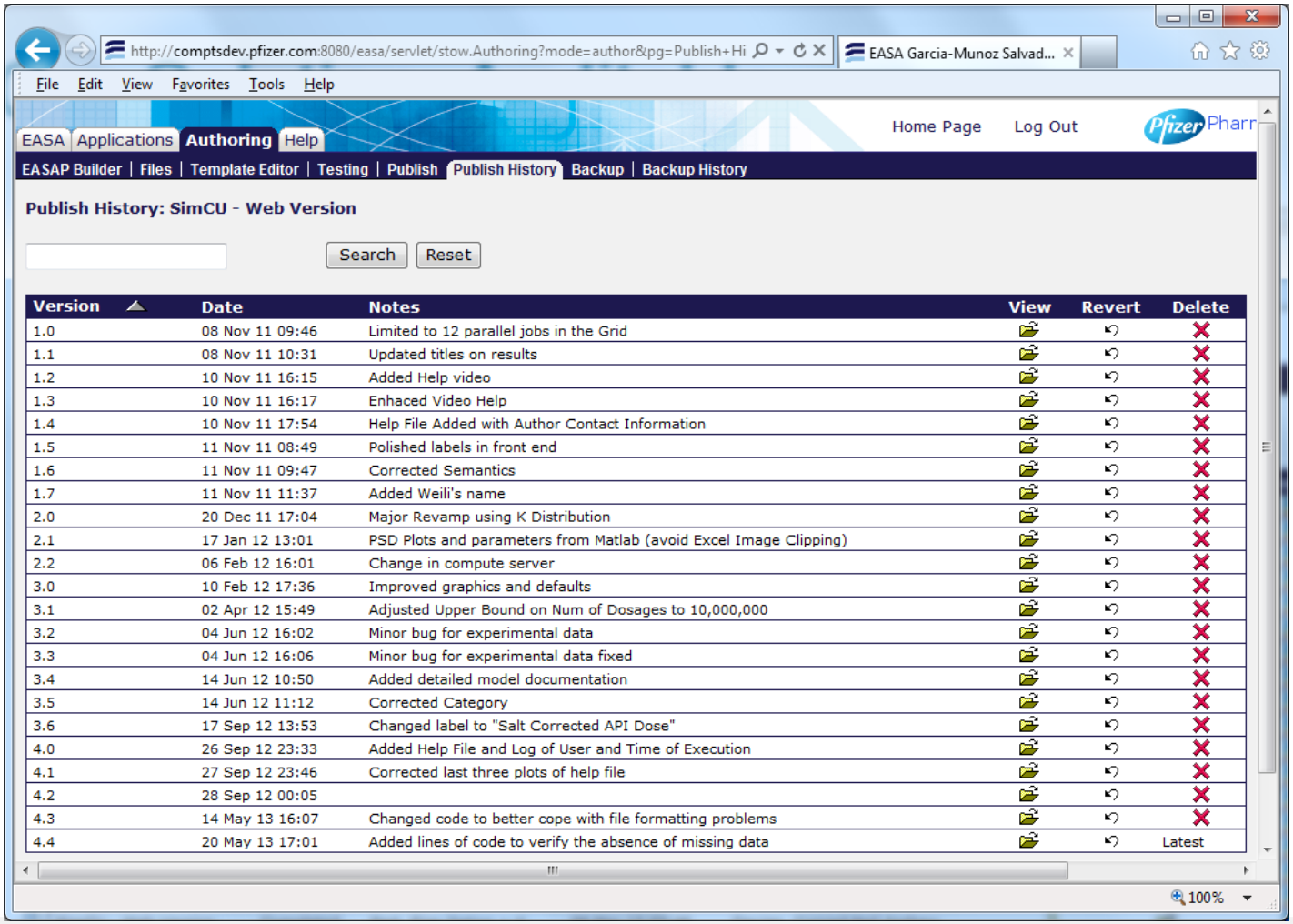


Application	Status	Notes	Date	User	Files	Inputs	Results
Integrated Genome-Wide Association Tool (InGWA)	Completed		03 Sep 13 12:12	Namburu, Lavanya Janaki Nivas			
Integrated Genome-Wide Association Tool (InGWA)	Completed		03 Sep 13 12:12	Namburu, Lavanya Janaki Nivas			
Integrated Genome-Wide Association Tool (InGWA)	Extracting available SNPs and P-value sources		03 Sep 13 12:12	Namburu, Lavanya Janaki Nivas			
Integrated Genome-Wide Association Tool (InGWA)	Extracting available SNPs and P-value sources		03 Sep 13 12:12	Namburu, Lavanya Janaki Nivas			
Blend Prediction Model - V 2.0	Completed	Group Name: **PF-05175157**	03 Sep 13 12:01	Namburu, Lavanya Janaki Nivas			
Lipid Data Quantification Tool (LDQT)	Completed		27 Aug 13 14:52	Lanba, Adhiraj			
Lipid Data Quantification Tool (LDQT)	Completed		27 Aug 13 14:08	Lanba, Adhiraj			
Integrated Genome-Wide Association Tool (InGWA)	Completed		27 Aug 13 14:05	Namburu, Lavanya Janaki Nivas			
Integrated Genome-Wide Association Tool (InGWA)	Completed		27 Aug 13 14:05	Namburu, Lavanya Janaki Nivas			
Lipid Data Quantification Tool (LDQT)	Completed		27 Aug 13 09:40	Lanba, Adhiraj			
Lipid Data Quantification Tool (LDQT)	Completed		27 Aug 13 09:37	Lanba, Adhiraj			

Our Solution: Tailor Fit

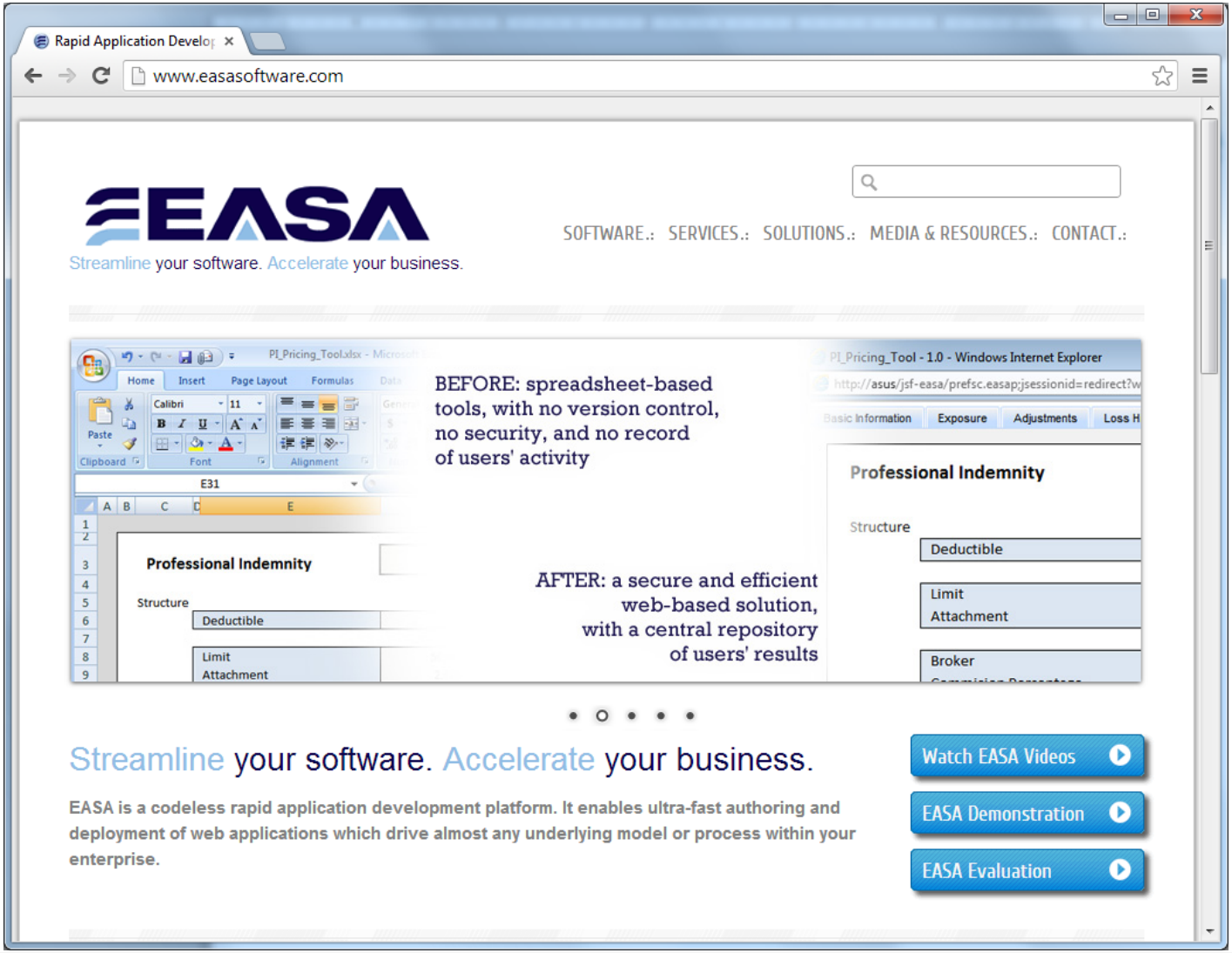


Our Solution: Central Version Control



Version	Date	Notes	View	Revert	Delete
1.0	08 Nov 11 09:46	Limited to 12 parallel jobs in the Grid			
1.1	08 Nov 11 10:31	Updated titles on results			
1.2	10 Nov 11 16:15	Added Help video			
1.3	10 Nov 11 16:17	Enhaced Video Help			
1.4	10 Nov 11 17:54	Help File Added with Author Contact Information			
1.5	11 Nov 11 08:49	Polished labels in front end			
1.6	11 Nov 11 09:47	Corrected Semantics			
1.7	11 Nov 11 11:37	Added Weili's name			
2.0	20 Dec 11 17:04	Major Revamp using K Distribution			
2.1	17 Jan 12 13:01	PSD Plots and parameters from Matlab (avoid Excel Image Clipping)			
2.2	06 Feb 12 16:01	Change in compute server			
3.0	10 Feb 12 17:36	Improved graphics and defaults			
3.1	02 Apr 12 15:49	Adjusted Upper Bound on Num of Dosages to 10,000,000			
3.2	04 Jun 12 16:02	Minor bug for experimental data			
3.3	04 Jun 12 16:06	Minor bug for experimental data fixed			
3.4	14 Jun 12 10:50	Added detailed model documentation			
3.5	14 Jun 12 11:12	Corrected Category			
3.6	17 Sep 12 13:53	Changed label to "Salt Corrected API Dose"			
4.0	26 Sep 12 23:33	Added Help File and Log of User and Time of Execution			
4.1	27 Sep 12 23:46	Corrected last three plots of help file			
4.2	28 Sep 12 00:05				
4.3	14 May 13 16:07	Changed code to better cope with file formatting problems			
4.4	20 May 13 17:01	Added lines of code to verify the absence of missing data			Latest

Our Solution: Powered by EASA Software



EASA System Components

Compute Servers

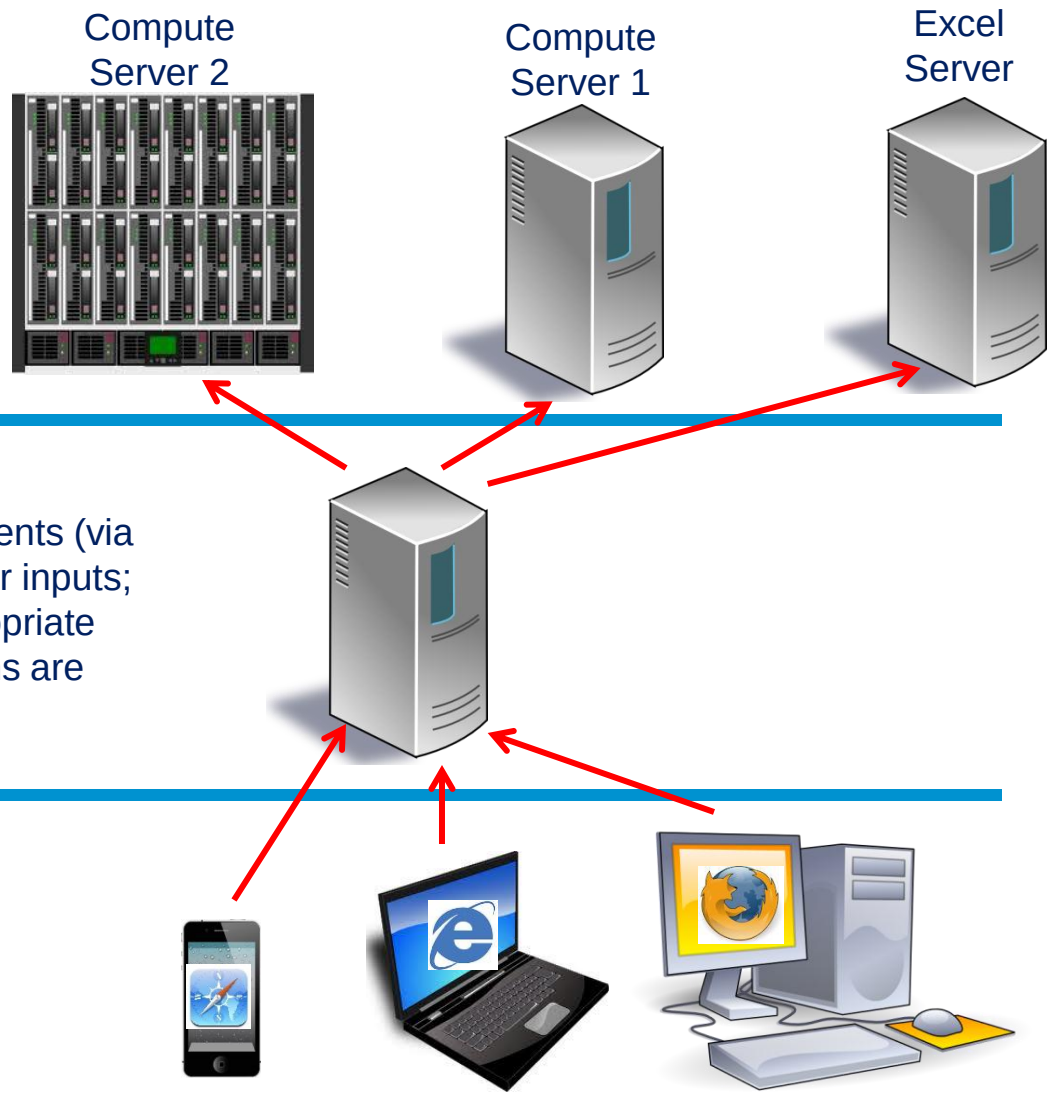
Computers with native application that communicate with the EASA server to perform calculations when requested. Excel servers are dedicated compute servers to execute Excel based applications.

EASA Server

Computer that communicates with clients (via http) to present model GUI and gather inputs; send out calculation jobs to the appropriate compute server; and once calculations are finished, present the results to client.

Clients

A web-capable device to access the EASA server



Lessons Learned I

- Needs in Pharmaceutical Sciences are **very** diverse
- “Model Maturity” is a relative term
- No such thing as one model for one problem
- It is relatively simple to identify candidate models for Web Deployment
  - Get’s hairy with data based models that are expected to change
- Users and SME’s are two different groups, routine application of these models is expected from Users
- Some models are to be kept in hands of the SME
- SME’s are building models in the best native environment without GUI-design concerns
- Model development is simplified since the programmatic are design for computer-to-computer interaction (not human-computer interaction)
  - Model updates are handled independently of GUI updates

Lessons Learned II

- Use the best modeling environment for the intended use
  - Why to use excel to solve an ODE system ?
- Computer-to-computer interactions are very easy to handle and make the model development VERY simple.
- This exercise helped us enforce a standard level of documentation
  - Help to understand the science behind the model
  - Help to know “Where do I click this thing”
  - How accurate is this model ? (what decision can I make with this?)
- It also has led us to better document ultimate model accuracy and business impact because:
- Such an easy access to models naturally drives the corporation to clearly state workflow impact for each model deployed

Lessons Learned III

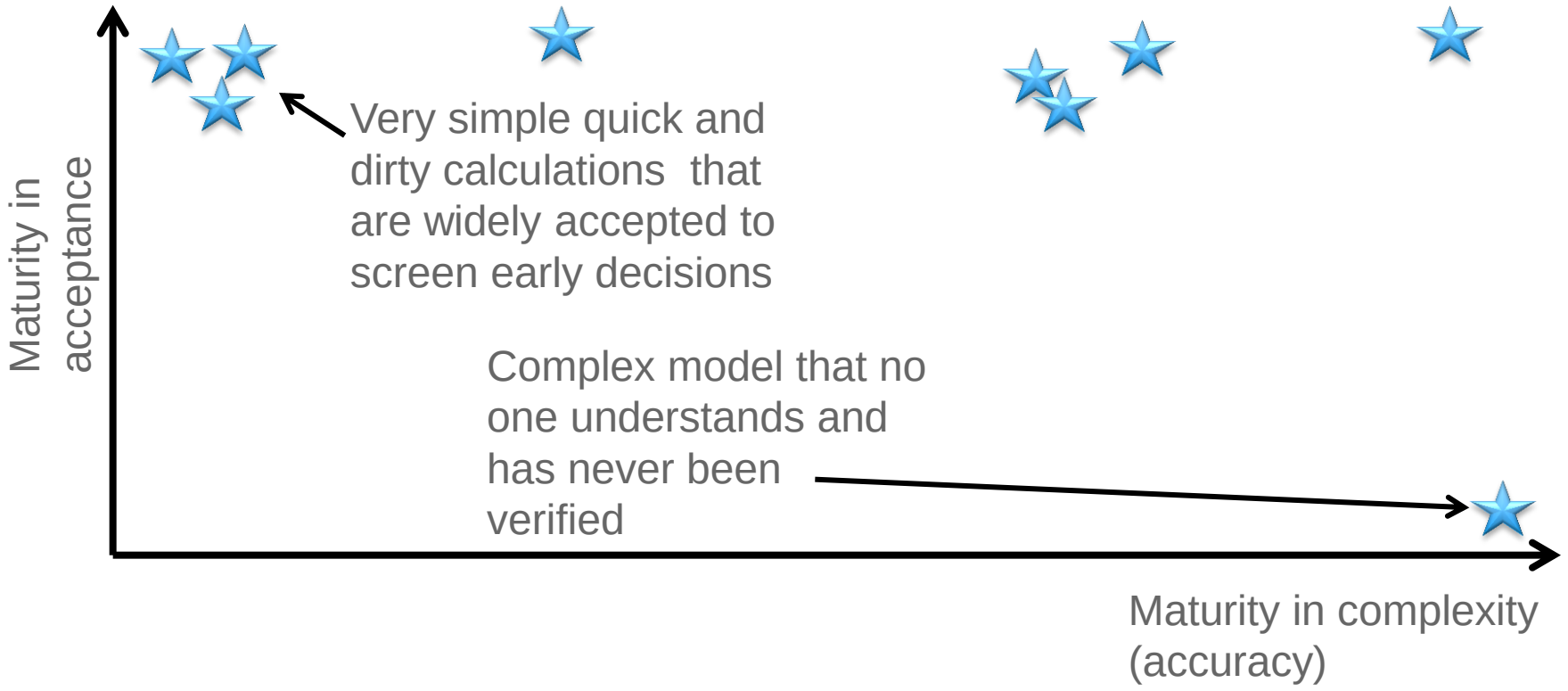
- Is the model accurate enough?
  - Interesting discussion…
  - To do what ?
  - To make what decision ?
  - This is a long-term discussion that requires the close dialog between SME and Users, focusing on how a model was effectively used to make an assertive decision.
    - How do we track that ??
- Benefit of the EASA platform is that multiple models can be published to make the same prediction with different levels of information needed and accuracy delivered.

Lessons Learned IV

- It is natural to have multiple models for the same purpose, which will become useful at different points in the development of a new product.
- Strong need to educate the user (or implement clever soft-interlocks) to prevent the wrong model from being used to support a given decision.
  - Use the GUI
  - Documentation
  - Classification in front end
- A tool such as this will impact culture of model-usage and at some point demand will require a critical mass of SME's to develop proper models to address business needs… if SME this support is not there... The system may collapse.

Lessons Learned V

- Metrics on model usage can become an important tool for resource management and allocation.
- Metrics will improve as models “mature”



Final Remarks

- EASA makes model deployment simple and robust for a corporation to engage the end user into a “self service” workflow when it comes to modeling technology.
- Licenses, Computer time and SME’s are optimally allocated.
- Auditability eases knowledge transfer and archival of results.
- Metrics can be drawn to support resource allocation decisions.
- IT and Business lines work closer in happy collaborative environment that leads to success.

Acknowledgements

- Our partner: EASA Software
  - Tim Valachovic
  - Seb Dewhurst
- Lavanya Namburu (Business Technology Support)
- EASA Authors (Pharm.Sci. SME's)
  - Fasheng Li
  - Vinicius Bonato
  - Bruno Hancock
  - Bill Ketterhagen
  - Natalie Culver
  - Eric Cordi
- Executive Sponsors
  - Keith Murphy (Pfizer Business Technology)
  - Debra Bremer (Pfizer Business Technology)

Questions?

Salvador Garcia-Munoz, Ph.D.