A Model-Driven Architecture For Unifying, Powering and Accelerating Drug Discovery
Technology-Driven Data

Biological Question

- Sequence
- Expression
- Structure
- Prot-prot interaction
- Polymorphism

Then
Technology-Driven Data

Then

Biological Question

Expression

prot-prot interaction

Sequence

Structure

Polymorphism

Now

High-Throughput Process

High-Throughput Process
Software in the Genomics Age

Presentation
Analysis
Access
Data

Secant Technologies: www.genomicsplatform.com | www.secant.com
Research Applications Are Developed And Run In “Silos”

- Researcher A
  - Sequence Application
  - Sequence Data

- Researcher B
  - Expression Application
  - Expression Data

- Researcher C
  - Pathways Application
  - Pathways Data

- Researcher D
  - Genetics Application
  - Genetics Data
The Problems With “Silos”

Secant Technologies: www.genomicsplatform.com | www.secant.com
The Problems With “Silos”

The Applications Are Not Integrated
The Problems With “Silos”

The Researchers Can’t Share Results

Secant Technologies:  www.genomicsplatform.com | www.secant.com
Results In A Broken Process

- Researcher A
  - Sequence Application
  - Sequence Data

- Researcher B
  - Expression Application
  - Expression Data

- Researcher C
  - Pathways Application
  - Pathways Data

- Researcher D
  - Genetics Application
  - Genetics Data

...a broken... …research... …process

Results in...
The Science Pain

- Too many “Targets” to integrate and maintain
- Wrong “Targets” put into screening
- Human Genes ≈ 30,000
  - Human expression and tissue specificity
  - Genetic variation (SNP’s)
- Cross reference other organismal genomes
- Personalized Medicine (predictive medicine vs. reactive)
The Business Pain

- $880M – Estimated cost to develop a novel drug
- Attrition rate of novel drugs ≅ 45% at clinical phase III
- $200M – Average amount that could be saved by eliminating one in 10 drug targets from research
- $300M – Savings estimated per novel drug if pharmas can properly implement and integrate genomics technologies and data along the way
- $1B – Estimated annual lost revenue for expired patent protection on 20mg Prozac blockbuster drug – Eli Lilly

Source: Boston Consulting Group
Introducing The GKP: Unifying The Process Of Discovery

The Genomics Knowledge Platform (GKP)
Integrated Data
Integrated Applications
Integrated Collaboration
Unified Research Process

Multiple Integrated Applications

Sequence Data
Expression Data
Pathways Data
Genetics Data
Viewing Data Through Science

The Genomics Knowledge Platform (GKP)
Integrated Data
Integrated Applications
Integrated Collaboration
Unified Research Process

Data is integrated to a Unified View (Model) of Biology

Researchers A, B, C, D

Multiple Integrated Applications

Sequence Data
Expression Data
Pathways Data
Genetics Data

Secant Technologies:  www.genomicsplatform.com | www.secant.com
Using A Model Of Science To Integrate Applications

Applications are integrated to a Unified View (Model) of Biology

The Genomics Knowledge Platform (GKP)
Integrated Data
Integrated Applications
Integrated Collaboration
Unified Research Process

Researcher A

Researcher B

Researcher C

Researcher D

Multiple Integrated Applications

Multiple Integrated Applications

Multiple Integrated Applications

Multiple Integrated Applications

Sequence Data
Expression Data
Pathways Data
Genetics Data

Secant Technologies:  www.genomicsplatform.com | www.secant.com
Enabling Collaboration Within And Between Organizations

The Genomics Knowledge Platform (GKP)
Integrated Data
Integrated Applications
Integrated Collaboration
Unified Research Process

Researchers can save their work and share it with others via the GKP server and the GKP’s application shell.
Unified Object Model
Analysis Layer

Alignment  
Feature  
Hit  
Annotation

Science  
Analysis Services
Services Layer

Robust, Scalable, Fault Tolerant

- Query, Save, Publish
- Result Sharing
- Data Loading, Versioning
- Workflow
- Install, License
- E-Commerce, AR, AP
- Collaboration, Security
- Object File System

Science
Analysis
Services
GKP Architecture

An Enterprise-Scale Platform
Distributed Architecture
Dynamic Load Balancing
Exponential Clustering
Parallel Processing

Search Engines
- Keyword Search
- Sequence Storage
- BLOB Storage

Extension Servers
- Job System API
- Job Controller
- Compute Farm API
- Farm Nodes

App Client
Web Client
Wireless Client
3rd Party Client

Base Module API Application
Base Module API Application
Base Module API Application

is Discovery Center
Application Services
GKP Object Model

ODBC

Data Integration

Data
Data
Data

Flat File

Secant Technologies: www.genomicsplatform.com | www.secant.com
GKP Architecture

Data Access

IBM’s Discovery Link

Data Integration Through Discovery Link Views

Discovery Link API

Driver

RDB

Driver

RDB

Driver

Flat File
GKP Architecture

Job Management

- Tool Registration
- Multi Job Task Definitions
- User Request Management
- Task Prioritization
- Parallelization of Single Tasks
- Support Results as Parameters to Other Tasks
- Allow Jobs to Register Jobs
- Plugable Distribution Logic
GKP Architecture

Application Framework

Data Access

Job Management

Unix / Linux Farm

Unix Farm API
Secant’s Model-Driven Infrastructure
ModelMethods™ Products

- Object Integrator for Java & C++
  - Object/Relational Mapping, Caching Technology
  - Modeling Tool Integrations
- Enterprise Server for Java & C++
  - Supports Java 2 Platform, Enterprise Edition
    - EJB, CORBA, Transactions, Security, Clustering
    - J2EE Licensee
  - Supports CORBA / C++
  - Modeling Tool Integrations
Model-Driven Approach

The OMG MDA Specification

Secant’s Model-Driven Infrastructure

Secant Technologies: www.genomicsplatform.com | www.secant.com
Development Methodologies Compared

Traditional Methods

- Design
- Code
- Test
- Code
- Test

TIME

Release 1 Deployed
Release 2 Deployed

MAINTENANCE

CHANGE 1

Secant’s Model Methods Process

- Design
- Model
- Code
- Test
- Code
- Test

NOTE:
Faster Development & Less Maintenance

Release 1 Deployed
Release 2 Deployed
Release 3 Deployed
Release 4 Deployed

MAINTENANCE

CHANGE 1
CHANGE 2
CHANGE 3

Auto Code Generation

Secant Technologies: www.genomicsplatform.com | www.secant.com
Model Method Development

Best-of-Breed Modeling Tool

1A Use As Is
1B Modify
1C Add New Objects

Object Definition Language Compiler

2 Automatic Code Generation

Remote Objects
Application Components
Class Code

Best-of-Breed IDE Tool

3 Add Application Logic
4 Deploy to Extreme Enterprise Server

Extreme Link Plug-In

Science
Analysis
Services

Secant Technologies: www.genomicsplatform.com | www.secant.com
Continuum in Value

Results in... Chemistry... researchers... process