New this Year

Future Work

genSpace: Community Driven Knowledge Sharing in geWorkbench

Gail Kaiser's Programming Systems Lab

enable (vt): to make possible, practical, or easy



http://www.psl.cs.columbia.edu/

Computer Science Department, Columbia University

June 1, 2012



Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

New this Year

Knowledge Sharing

- Scientists collaborating together in the same lab on the same project share:
 - Data: specimens, samples, materials, analyses
 - Tools: instruments, software, hardware
 - Knowledge: open discussion, whiteboard
- However, there are temporal (time) and physical (space) constraints
- This model does not scale to communities of scientists working on different projects but who could possibly learn from each other's expertise, experience, etc.



New this Year 00000

CSCW Approaches

- Most current generation Computer-Supported Cooperative Work systems enable data sharing and/or tool sharing (e.g., EMSL Collaboratories, UIUC BioCoRE, Monod)
- However, these systems support relatively limited knowledge sharing - how/when/where/why to use tools and data
- Knowledge sharing is partially enabled through labor intensive approaches: pubs, email lists, wikis, chat, shared display, etc.
 — may be outdated, requires active participation
- We seek to enable automatic knowledge sharing without requiring "extra work" by scientists



New this Year 00000

genSpace and geWorkbench

- As geWorkbench continually expands with new analysis and visualization modules, choosing the right tools to use and learning how to chain these tools in sequence (workflows) can be very daunting, especially to novice users
- genSpace aims to alleviate this problem by providing recommendations about what to do first and what to do next
- Records, aggregates and data mines geWorkbench users' activities with tools and tool sequences
- Users can ask questions such as what are the most frequently used tools and which workflows start with or contain a given tool, with the answers derived (and updated periodically) from the collective experience of geWorkbench users



Introduction to genSpace 000●	genSpace Features 00 0000	New this Year 00000	

Features

Our overall goal is to deliver community driven knowledge sharing, and to that end, we have implemented many features including:

- Tool and workflow recommendations
- Workflow sharing and visualization
- Community portal (website for tool and workflow comments)
- Social networking features (friends, networks, chat, profiles)
- Research notebook with automatic updates



	genSpace

genSpace Features ●0 ○○○○ New this Yea

Recommendations

Workflow Visualization

Viewing all workflows with a given tool





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Recommendations

Workflow Visualization

Results





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

genSpace Features ●0 ○○○○ New this Year

Recommendations

Workflow Visualization

Rating tools and identifying expert users

000			genSpa	ce				
genSpace Login	Workflow Visualization	Workflow Suggestions	Workflow Statistics	Social Center	Research	Notebook	Workflow Repository	
	select action	💌 Ano	va Analysis			Search	150 workflows found	
	Vora Analysii Vora Childi Vora Childi Vo	Clustering Constant Expert Constant Expert Real Net Vertice/add workflow Vertice/add work	page for Hierarchie User - USB Tobber Source - USB to Tob Source -	al Clustering () (sees.) () () () () () () () () () () () () ()				Ď
Status: Ready	Anova Anal	riii Brown Bradmin	,					,_

CS± @CU

Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Recommendations

Workflow Suggestions

Default view





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Recommendations

Workflow Suggestions

Can be focused by tool



Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Recommendations

Workflow Suggestions

Updates as user performs analyses

perSpace Login Workflow Visualization Workflow Suggestions Workflow	v Statistics Social Center Research Notebook Workflow Repository
Ally Detection Cell Piller	
ur current workflow: fy Detection Call Filter	Advanced suggestions Get suggestions for people who use these tools:
opperione for your head drags: or a straft with the has have used 32 times by get@gete users. The straft with the has have used 32 times by get@gete users. The straft with the has have used as the straft with the has have been about the straft with the straft with the has have about the straft with the	Any Detection Call Files Alware Analysis Much Z Analysis Call Analysis C
itus: Ready	



Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features ○○ ●○○○	New this Year 00000	
Community based knowledge			

Workflow Repository

Before accepting a shared workflow

product comp Workflow Subjects Social Cetter Workflow Repository ** Trendsity's Audyin Catatring Example Social Cetter Workflow Repository ** Trendsity's Audyin Catatring Example Social Cetter Workflow Repository ** Trendsity's Audyin Catatring Example Social Cetter Workflow Repository ** Trendsity's Audyin Catatring Example Social Cetter Workflow Repository ** Defense Softwared Social Cetter Workflow Repository ** Defense Softwared Workflow Repository Social Cetter ** Defense Softwared ** Social Cetter Workflow Repository ** Defense Softwared ** Workflow Repository ** Social Cetter ** Trendsity Audyin ** Social Cetter ** Social Cetter ** Defense Softwared ** Workflow Repository ** Social Cetter ** Social Cetter ** Social Cetter ** Social Cetter ** Social Cetter ** Social Cetter ** Social Cetter	000			genSpace			
Wetrostov Tetostov Districtor Bist Tativiti AdXU/Matchen AdXU/Matchen Wetrflow Information	genSpace Login Wo	orkflow Visualization	Real Time Workflow Suggestion	Workflow Statistics	Social Center	Workflow Repository	
Detex Selected Workflow Index Name: Uoder Date Workflow Commercs: Workflow Details Workflow Commercs: Workflow Details Workflow Information Workflow Information	Tuesday's Ana Clustering Execution Basic Tutorial ANCVA Work! Swapned's Wo	tilvis Innole Iow orkflow				<u></u>	2
Long L. Wed M. > Lonk A pin Tar M Send Import Export Publish Refresh Workflow Information	Delete Selecter	d Date Workfli	ow Comments Workflow Details	1			
Add Dekte	Add De	elete	low Information	end Import	Export Pub	lish Refresh	

Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features ○○ ●○○○	New this Year 00000	
Community based knowledge			

Workflow Repository

After accepting a shared workflow

000			genSpace			
genSpace Login	Workflow Visualization	Real Time Workflow Suggestion	Workflow Statistics	Social Center	Workflow Repository	
General Construction General Construc	Analysis Analysis Example rat Sriflow Workflow				<u></u>	
Delete Sele	rcted 🗡					
Name User	Date	ow Comments Workflow Datails	1	_		
Some npun Look a jon	Wed M Tue M Delare	low Information	end Import	Export Pub	lish Refresh	*
Statur: Ready						~



Computer Science Department, Columbia University

Introduction to genSpace	genSpace Features ○○ ●○○○	New this Year 00000	
Community based knowledge			

Workflow Repository

Viewing a workflow in the repository





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace	genSpace Features	
	00 0 0 00	

Community based knowledge

Workflow Statistics

gamma participation Water bargenom Weater bargenom Weater bargenom Weater bargenom Weater bargenom 	00	gerSpace
Most Popular Tools Most Popular Cools at Start of Workflow 9: A way	genSpace Login Wo	orkflow Visualization Workflow Suggestions Workflow Statistics Social Center Research Notebook Workflow Repository
Instantian AddArd Acquing, AddArd Acquing 9 AddArd Acquing, AddArd Acq		Most Popular Tools Nost Popular Tools at and Workflow 1. AMCUT Avainaming 1. Hearchead Learning 1. Tort E avainaming 1. Hearchead Learning 1. MADUT Avainaming 1. MAUCH Avainaming 1. MAUCH Avainaming 1. MAUCH Avainaming 1. Kon Avainaming 1. MAUCH Avainaming 1. Kon Avainaming 1. MAUCH Avainaming 1. Kon Avainaming 1. SAM Avainaming 1.
1. AACIA Kangia, HAKOA Kangia 1. August Managa		Most Popular Workflows
Individual Tool Statistics Aeroa Analysis Tool laster not: 513 Tool laster not: 514 Tool laster not: 514		AACHI Avalysis, AACHI Avalysis AACHI Avalysis AACHI Avalysis AACHI Avalysis Article Avalysis Test Avalysis Mak Avalysis Mak Avalysis AACHI Avalysis, Test Avalysis AACHI Avalysis, Annu Avalysis, MACHI Avalysis, AACHI Avalysis AACHI Avalysis, Annu Avalysis, AACHI Avalysis, Annu Avalysis, AACHI Avalysis, AACHI Avalysis Sol Avalysis
Const Analysis The Taral scape rate at some of a software (*14) The Issue rate at some of a software (*14) The Issue rate at some of a software (*14) The most popular bod used before this bod instanticepang Concer Normalizer The most popular bod used before this bod instanticepang Conce Normalizer		Individual Tool Statistics
Trace using one: 2715 Trace using one at stand or workflow: 741 The most popular tool used were to the tool Staffast Analysis. The most popular tool used keller the tool Husself caping Creek Normalizer The most popular tool used keller the tool Husself caping Creek Normalizer atter. Ready		Anova Anabsis
Ratus: Ready		Tool usage rate 27.15 The inset parallel to disaff net to the look Status Autyo's The most parallel to disaff net to the look Status Autyo's The most parallel to disaff net to look Status Autyo's The most parallel to disaff net look Status Autyo's The most parallel to disaff net look Status Autyo's
	tatus: Ready	



Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features ○○ ○○●○	New this Year 00000	
Community based knowledge			

Viewing friends





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features ○○ ○○●○	New this Year 00000	
Community based knowledge			

Viewing profiles

00			genSpa	ce			
genSpace Lo	gin Workflow Visualization	Workflow Suggestions	Workflow Statistics	Social Center	Research Notebook	Workflow Repository	
GenSpace					54	arch	Go
	tudo Tõhhon (12102)'s gonfn	aco profilo				
ty Profile	Udo Többen (ut2102)	atz 102) s gensp	ace prome				
V Networks	at C282						
l <u>y Friends</u> had	Research Interests: cEF1A						
Requests	Contact information:						
iew Requests	Phone: 2128514664 Email: ut2102@columbia.ed						
	Mailing Address: 1130 St. Nicholas Avenue						
	Room 808H New York NY, 10032						
	(Udo is a friend) Remove friend						
Status: Ready							



Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features ○○ ○○●○	New this Year 00000	
Community based knowledge			

Users by network and chat





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features ○○ ○○●○	New this Year 00000	
Community based knowledge			

Friend and network requests



CS∰ @CU

Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features ○○ ○○●○	New this Year 00000	
Community based knowledge			

Privacy settings





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features ○○ ○○○●	New this Year 00000	
Community based knowledge			

Community portal

Tool list





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features ○○ ○○○●	New this Year 00000	
Community based knowledge			

Community portal

Tool detail





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features ○○ ○○○●	New this Year 00000	
Community based knowledge			

Community portal

Workflow detail

	na anto incontinuo (10)		
Workflow Detail	(+ [cga	
	genSpace		
	Senspace .	Ack Tool Search: Q	
	Abert 1005 Septert		
	geWorkbench workflows		
	View Workflow:	Login	
		Logged in as Jonathan Bell.	
	Steps: MINDY Analysis, MINDY Analysis, T Test Analysis, SON Analysis, T Test Analysis, T Test Analysis, WINDY Analysis, T Test Analysis comments	AND COLLECT	
	There are no comments for this workflow.	Rating	
		BATE THIS WORKFLOW:	
	Submit a comment:	Rate this: Good	
		submit	
	Eshell Dura		
	amount down A		
	ry 2010 - Research Columbia University Researching Sectores Lab. Daries by Node/Del	Turne - Ener (SS Tarrel stor	
	o 2010 - Present Columbia University Programming Systems Lab. Design by HodeThin	ty/Three + Free CSS Templates	
	o 2010 - Present Columbia University Programming Systems Lab. Design by HodeThit	tyThree - Free CSS Templates	
	o 3310 - Present Columbia University Programming Systems Lab. Design by HodeThit	ty/Three + Free CSS Templates	
	o 2016 - Present Galandra University <u>Programming System Lab</u> , Design by <u>HoofThe</u>	tyThnee + Free CSS Templates	
	e 2014 - Presen Gaundia University <u>Programming Sprawn Lab</u> . Design by <u>HostPhr</u>	n/Three - Free CSS Templates	
	e 2019 - Annae Galandia (antonia) <mark>Anganeting Samera (an</mark> Bollyn in <u>Haddha</u>	ryThree - Free CSS Templanes	
	E-1014 - Annuel Constant Internets (<u>Respecting Lemma Lab.</u> Internet in <u>ResPER</u>)	yThese - Free CSS Templates	
	e 2015 - Anneen Calanda Lanneau, <u>Ingerenerg Lanneau, da</u> Annye na <u>Handha</u>	tyfhane - Free CSS Templates	
	2.2011 - Jones Calento Donesto <u>Suggerenz Jones</u> do Deny in <u>Bachte</u>	ryThere + Erec CSS Transformer	
	e 2015 - Anneer Galantin Ganerin (Ingenering <mark>Joseph und Annee in Handhin</mark>	yihane - fine CSI Temaining	
	2.2011 - Jones Calento Danesto <u>Suggerenz James Las</u> Integri o <u>Barthe</u>	yfhae - fae CSI Tenylane	
	e 2011 - Annee Calanda Canneas, <u>Ingerorrang Joseph (a</u> Indep in <u>Bachte</u>	n Than - Fare Cill Transform	

Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

Introduction to genSpace 0000	genSpace Features 00 0000	New this Year ●0000	

This year's focus was on socially-oriented recommendations and collaboration tools.

- "People like you"
- Collaborative filtering and diversity in recommendations
- Research notebook

Additional notable tasks completed:

- Re-engineered genSpace recommendation platform to use Apache Mahout (scalable open source machine learning library)
- Developed prototype web-based genSpace user interface to integrate with the in-development web-based geWorkbench



	genSpace

New this Year

"People like you"

Based on social connections and geWorkbench usage

annSaura Logia Ù Workflor	Workflow Sugar	9 actions T Work	flow Statistics	Social Center	Rerearch Notebook	Workflow Reportion
genspace Loginworknow	v visualization worknow sogg	esuons work	now statistics	Social Center	Kesearch Notebook	worknow Kepository
	Data Visibility Genera	I Profile				
	First Name	J	onathan			
	Last Name	E	ell			
	Lab Affiliation "	F	SL			
	Email Address					
	Phone					
	Address 1					
	Address 2				logout	
	City					
	State					
	ZIP Code					
	Save					
Your Recom	mended Workflows		F	eople Like Y	ou	Filter to My Networks
ARACNE Analysis, ARA Hierarchical Clustering Hierarchical Clustering MRA Analysis Anova Analysis	CNE Analysis , Hierarchical Clustering , T Test Analysis	1. sw 2. lee 3. ch 4. ka 5. Ar	rap – Swapneeh onw – Leon Wu ris – Chris Mur iser – Gail Kais ionymous	i Sheth phy er		
tus: Ready						

CS± @CU

Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

New this Year

Recommendations based on collaboration networks

- genSpace tool recommendations now take into account a user's collaboration networks (including implicit "people like you" networks)
- This creates an issue of diversity:
 - Personalized recommendations can lead to a "bubble"
 - For example, if all members of a lab's recommendations are based heavily on their lab mates, there will be no exposure to new ideas
 - genSpace solution: mix in results from other (visible) networks automatically, and allow users to specifically request more diverse results



New this Year 000●0

Research Notebook

- Permanent archive of geWorkbench experiments
- Automatically logs every analysis, including dataset and parameters
- Annotations can optionally be added to any analysis
- History is fully searchable



New this Year 0000●

Research Notebook

Viewing the notebook





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

New this Year 0000●

Research Notebook

Editing a notebook entry

000			genSpa	ice			
genSpace Login Workfl	ow Visualization	Workflow Suggestions	Workflow Statistics	Social Center	Research Notebook	Workflow Repository	
ilter:							
Enter your search query he	re or use the drop	odown below					Search
							•
ly Log:							
Now editing (Esc to canc	el)						
Affy Detection Ca Dataset: Bcell-100	ll Filter at 3/5/. .exp	12 11:59 AM					
This didn't do what I thou	ght it would. Nee	d to try again.					
		_ v i	w Parameters	Cancel Save]		
Affy Detection Call Filte Dataset: 8cell-100.exp	r at 5/18/12 1	0:37 AM					
Quantile Normalizer at Dataset: web100.exp Lorem ipsum dolor sit amet aliquam rhoncus dignissim.	S/10/12 2:22 i	PM Hiscing elit. Donec tincidu	nt gravida egestas. Ir	nteger erat leo, e	lementum sit amet hen	drerit sed, aliquet non	sapien. Fusce
Hierarchical Clustering Dataset: brain mas5_170	at 5/10/12 2:1 filtered.exp	9 PM					
Duis convallis ligula at eros lacinia nunc egestas viverra	faucibus ornare. I Nam condimenti	Pellentesque auctor fring um volutpat sem nec euis	illa magna quis matti: imod.	s. Vestibulum ne	c tellus non leo suscipit	vestibulum. Suspendis	se nec neque
Hierarchical Clustering Dataset: aTestDataSet.ex	at 5/10/12 2:1 p	1 PM					
Praesent urna ante, tristique montes, nascetur ridiculus r	e eu ultrices vel, a nus.	uctor vel massa. Aliquar	n in ipsum in velit ele	ifend laoreet in o	get mi. Cum sociis nato	ique penatibus et mag	nis dis parturient
Standard Deviation Filte Dataset: brain_mas5_176 THIS WAS MY TEST	er at 3/8/12 1: _filtered.exp	55 PM					
Marker Centering Norm	alizer at 3/8/1	2 1:54 PM					



Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

New this Year 0000●

Research Notebook

Viewing parameters from an analysis





Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

New this Year 0000●

Research Notebook

Filtering view by tool

000			genSpa	ce			
gerSpace Login	Workflow Visualization	Workflow Suggestions	Workflow Statistics	Social Center	Research Notebook	Workflow Repository	
lter:							
							Search
Anova Analysis							•] — •
lv Log:							-
Anova Analysis Dataset: 21161 n	at 2/7/12 5:36 PM eu10.gpv						
Anova Analysis Dataset: 21161 n	at 2/7/12 5:36 PM eu10.gpr						
Anova Analysis Dataset: 21161 n	at 2/7/12 5:35 PM eu10.gpr]
Anova Analysis Dataset: 21161 n	at 1/10/12 4:52 PM euI0.gpr						
Anova Analysis Dataset: 21161 n	at 8/3/11 7:15 PM eu10.gpr						_
Anova Analysis Dataset: 21161 n	at 8/3/11 7:15 PM eu10.gpr						
Anova Analysis Dataset: 21161 n	at 7/26/11 3:01 PM eu10.gpr						_
Anova Analysis Dataset: 21161 n	at 7/26/11 3:00 PM eu10.gpr						
Anova Analysis Dataset: 21161 n	at 7/26/11 2:59 PM eu10.gpr						
Status: Ready							



Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University

New this Yea 00000

Future Work

- Expand research notebook feature to allow groups of researchers to share notes
- Fully develop a web-based genSpace to complement the web-based geWorkbench
- Full support for workflows that include aborted and long-duration tools in geWorkbench



New this Year

Future Work

genSpace: Community Driven Knowledge Sharing in geWorkbench

Gail Kaiser's Programming Systems Lab

enable (vt): to make possible, practical, or easy



http://www.psl.cs.columbia.edu/

Computer Science Department, Columbia University

June 1, 2012



Gail Kaiser's Programming Systems Lab

Computer Science Department, Columbia University