Cyberintegrator Manual

Cyberintegrator (CI) was created at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign. We would like to acknowledge multiple funding agencies for the support including NCSA, NSF, ONR TRECC and NARA. The main creators of CyberIntegrator are Rob Kooper, Luigi Marini and Peter Bajcsy, and the support from Jim Myers. This document represents a current description of our on-going research and development of Cyberintegrator and hence this document is updated on a regular basis.

Revision History

Revision	Date	Author	Notes
0.1	4/04/2006	PB, RK,LM	Initial version of the document
2.0	10/05/2007	PB, RK,LM	Major upgrade
3.0	07/09/2008	RB	Third Revision

Cyberintegrator Manual	1
Revision History	3
Revision History	3
Introduction	5
Introduction 1.1 Main Cyberintegrator Functionality 1.2 Cyberintegrator User Interface 1.3 Preferences	5 6 7 12
Chapter 2 Getting Started 2.1 New Workflow and Login	20 20 22
2.2 Importing and Exporting Data Existing Tools and Worktlows	••••••
2.2 Importing and Exporting Data, Existing Tools and Workflows2.3 Viewing and Annotating Data, Tools and Workflows2.4 Creating a Tool	30 40
 2.2 Importing and Exporting Data, Existing Tools and Workflows 2.3 Viewing and Annotating Data, Tools and Workflows 2.4 Creating a Tool Appendix A Software License 	30 40 52
 2.2 Importing and Exporting Data, Existing Tools and Workflows 2.3 Viewing and Annotating Data, Tools and Workflows 2.4 Creating a Tool Appendix A Software License 	
 2.2 Importing and Exporting Data, Existing Tools and Workflows 2.3 Viewing and Annotating Data, Tools and Workflows 2.4 Creating a Tool Appendix A Software License Appendix A Software License Appendix B Known Issues 	

Introduction

As the amount of observation data captured everyday increases, running scientific workflows will soon become a fundamental step of scientific inquiry. Current scientific workflow systems offer ways to link together data, software and computational resources, but often accomplish this by requiring a deep understanding of the system with a steep learning curve. Thus, there is a need to lower user adoption barriers for workflow systems and improve the plug-and-play functionality of these systems which drove our development of Cyberintegrator¹.

Cyberintegrator is a user friendly editor to several middleware software components that (1) enable users to easily include tools and data sets into a software/data unifying environment, (2) annotate data, tools and workflows with metadata, (3) visualize data and metadata, (4) share data and tools using local and remote context repository, (5) execute step-by-step workflows during scientific explorations, and (6) gather provenance information about tool executions and data creations.

The middleware software components interfaced by Cyberintegrator editor are (a) multiple plug-in codes called executors for running external tools, Matlab scripts and java codes, (b) Tupelo data and metadata archiving system based on the Resource Description Framework (RDF) metadata model, and (c) portal systems (currently Cybercollaboratory based on Liferay portal technology).

The Cyberintegrator editor is built on top of the eclipse rich client platform (RCP) to achieve its full functionality of re-configurable user interface and plug-and-play architecture critical for easy deployment of the Cyberintegrator application. The ultimate goals of Cyberintegrator are to support scientific discoveries and to provide means for conducting research more efficiently.

Background

The functionality of Cyberintegrator that allows linking software tools is similar to a workflow system. Current paradigms for workflow creation focus on the visual programming using a graph based metaphor. This can be a powerful metaphor in the hands of expert users, but can become daunting when graphs become large, the steps in the graph include engineering level steps such as loading and visualizing data, and the users are not very familiar with all the possible tools available. Cyberintegrator presents a different method of workflow creation that co-exists with the standard graph based editors. The method builds on exploratory interface using a macro-recording style, and focuses on the data being analyzed during the step by step creation of the workflow.

Instead of storing data in system specific data structures, the use of more flexible open standards that are platform independent would create systems that are easier to extend and that

¹ More information about Cyberintegrator project can be found at <u>http://isda.ncsa.uiuc.edu/software.html</u>

provide a simple interface for external applications to query and analyze the data and metadata produced. We have implemented Cyberintegrator that stores workflows and related metadata using the Resource Description Framework (RDF) metadata model. It is build on top of the Tupelo, Sesame and WebDAV technologies. Tupelo is a data and metadata archiving system. Sesame is a computing infrastructure designed to support secure interactions among the client, server, and database tiers, and application modules that carry out specific tasks. WebDAV is Web-based Distributed Authoring and Versioning software for editing and managing files collaboratively on remote <u>World Wide Web</u> servers. Cyberintegrator connects to content repositories (local or remote), where users can easily share data, workflows, algorithms and annotations.

In comparison with Cyberintegrator version 1, this version 2 is built on top of a Rich Client Platform (RCP) technology. A Rich Client Platform (RCP) could be viewed as a computer program designed for the creation of applications written in Java and built on the Eclipse architecture. Eclipse is an Integrated Development Environment (IDE) that is open source and developed by the on-going Eclipse project members² and by the Apache Software Foundation.³ Applications written using RCP are portable and the plug-ins can be reused and mixed.

From this perspective, the Cyberintegrator version 1 was lacking and hence the redesign leading to version 2. The Cyberintegrator version 2 can now support constantly changing requirements of a typical exploratory science. The RCP platform is also ideal for (1) accommodating user-driven customizations of graphical user interfaces (GUI) of the workflow editor, (2) adding new features to the workflow environments in run-time and (3) deploying Cyberintegrator on multiple platforms and with modules/components custom-selected by the end users of the workflow system.

1.1 Main Cyberintegrator Functionality

The key functionalities of the current implementation of Cyberintegrator can be summarized as follows.

(1) It enables users to easily include tools and data sets into a software/data unifying environment.

(2) Cyberintegrator comes with a set of functionality for visualizing text/ASCII and image data, as well as metadata and annotations about data.

(3) It enables users to annotate data, tools and workflows with metadata.

(4) Cyberintegrator allows sharing data and tools using local and remote context repository,

(5) It enables scientific explorations by step-by-step execution of workflows.

² http://www.eclipse.org/membership/exploreMembership.php

³ http://www.apache.org/

(6) Cyberintegrator gathers metadata information about imported tools and data, as well as about tool executions and data creations.

1.2 Cyberintegrator User Interface

The graphical user interface (GUI) of Cyberintegrator consists of views (or windows displaying certain information about data, tools or status) and perspectives (a collection of views). One perspective (GUI configuration) of the Cyberintegrator editor is shown in Figure 1.

* •	Cyberintegrator				و الماري	a x
Eile Edit Window Help						
0				1		39
🔒 Data 🕄 🔍 🔍 🖓	Document: Wind Speed Document: Win				â 🔨	- 0
type filter text	20060510T12:00:00 153 20060510T12:06:00 NA					Ĥ
Wind Direction (DONE)	20060510712:12:00 158					
🗟 Wind Speed [DONE]	20060510112:18:00 NA 20060510712:24:00 NA					
🔒 retval [DONE]	20060510T12:30:00 153 20060510T12:36:00 NA					
	20060510T12:42:00 NA 20060510T12:48:00 NA					
	20060510712:54:00 158					
	20060510T01:06:00 NA					
	20060510T01:12:00 NA 20060510T01:18:00 NA					
	20060510T01:24:00 NA 20060510T01:30:00 156					
	20060510T01:36:00 NA 20060510T01:42:00 172					
	20060510T01:48:00 NA					
	20060510T01:54:00 NA 20060510T02:00:00 157					
	20060510T02:06:00 NA 20060510T02:12:00 160					
	20060510T02:18:00 NA					
	20060510T02:30:00 159					
🕿 Tools 🕅 🦉 🍟 🗖	20060510102:36:00 NA 20060510T02:42:00 163					
type filter text	20060510T02:48:00 NA 20060510T02:54:00 160					
🔐 TCOON data (v1)	20060510T03:00:00 160 20060510T03:06:00 NA					
🔗 Transform TCOON wind data (v1)	20060510703:12:00 158					
	20060510T03:24:00 NA					
	20060510T03:30:00 154 20060510T03:36:00 NA					
	20060510T03:42:00 157 20060510T03:48:00 NA					
	20060510T03:54:00 163					
	20060510104:06:00 NA					
	20060510T04:12:00 NA Preview Annotations					Ŧ
(Constant on M					
worknow or april 2 loor Execution	Servir cog to	phus is	Data.	10 0, 08 • 1		
Message	louidata (103303h3, 1-32, 14bd, br28, 3h0a58d8-3a2 as Booste bariate and or 1710/63800 h	Plug-In	Date 2008-07-00 15-14-17 105			
Could not load latest version of context from	iow/data/40320303-183c-4000-0c28-300e38d882e2 as /tmp/cyderincegrator/1749403809.0 w disk	edu uiuc nosa cyberint	2008-07-09 15:14:17:105			
NLS missing message: ContextSetup NoDe	erby in: edu.uiuc.ncsa.cyberintegrator.ui.messages	org.eclipse.osgi	2008-07-09 15:10:18.368			
1 Converted 28 triples, 0 documents, 0 step of	data, 0 tool data, 0 parameters, and 2 blobs to new format.	edu.uiuc.ncsa.cyberint	2008-07-09 15:10:18.345			
i Could not load latest version of context from	n disk.	edu.uiuc.ncsa.cyberint	2008-07-02 13:42:35.952			
MLS missing message: ContextSetup_NoDe	erby in: edu.uiuc.ncsa.cyberintegrator.ui.messages	org.eclipse.osgi	2008-07-02 13:42:35.520			
i Converted 28 triples, 0 documents, 0 step o	data, 0 tool data, 0 parameters, and 2 blobs to new format.	edu.uiuc.ncsa.cyberint	2008-07-02 13:42:35.499			
Problems saving preferences.		org.eclipse.core.runtin	2008-07-02 13:33:02.295			
Exception saving preferences to: /home/ralp	oh/cyberintegrator/Cyberintegrator/workspace/.metadata/.plugins/org.eclipse.core.runtime/.	org.eclipse.equinox.pr	2008-07-02 13:33:02.288			
t recored http://cet.ncsat.uuc.edu/2007/work	www.uarapezizoworoodi.edudeazisi-orvescosderr as /rmp/cybenniegrator/retvalixt	eou.duc.ncsa.cybennt	2000-07-02 13:31:32.307			•
				8	M of 10M	Û

Figure 1. A set of views creating a perspective. The views display imported data sets (left upper corner), tools (left middle), a set views corresponding to the workflow, workflow graph, step creator and error log (bottom) and metadata information view (right large pane).

Views and perspectives can be manipulated (create, open, close, save, reset) from the main menu (see Figure 2). Each view could be maximized by double clicking on the view tab and restored by double-clicking the view tab again. Views can be reconfigured by dragging and resizing each view window. Several example perspectives and views are illustrated in Figure 4, Figure 5 and Figure 6.

When a set of views was customized then it could be saved by selecting "main->window->save perspective as". The same perspective could be opened by selecting one of the saved perspectives in the dialog launched by "main->window->open perspective->other" (see Figure 3) There are 3 preset perspectives (Cyberintegrator, Repository Exploration, and Workflow Execution) that come standard with the Cyberintegrator package, as shown below.

Bit Bit groten jub/ Bit State St				Cyberintegrator			+0.81
Open Progenting Open Progenting <th>Eile Edit Window Help</th> <th></th> <th></th> <th>0.0000000000000000000000000000000000000</th> <th></th> <th></th> <th></th>	Eile Edit Window Help			0.0000000000000000000000000000000000000			
Service Tepstop (p) model Tepstop (p) model Service Tepstop (p) model Tepstop (p) mod	Open Perspective	" Cyberintegrator					
Date 3 per se regregator 6 Oper 1 fill words execution Oper 1 fill words execution Oper 1 fill words Oper 1 fill words Oper 2 fill words	Show View +	Pepository Exploration					<u>.</u>
Up the Number of Decision of Decisi	Data 2 Save Perspective &s	C Workflow Execution	nt: Wind Speed 23	Annotations			() × = (
Busk Bright Angenetics Disk Bright Angenetics Disk Bright Angeneti	type filter Beset Perspective	Qther	T12:00:00 153				1
Care of language Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob Prob </td <td>Glose Perspective</td> <td>20060510</td> <td>0712-12:00 158</td> <td></td> <td></td> <td></td> <td></td>	Glose Perspective	20060510	0712-12:00 158				
Image for the measure of the set of	Close All Perspectives	20060510	0T12:24:00 NA				
Humin TCOSH wind data (1) Society 112:2:2:00 HA Society 112:2:00 HA Society 112:2:	B rety Dreferences	2006051	0712:30:00 153				
	 Brefs 31 Open Mart text ∰ toors data (v1) ∰ transform TCODE und data (200651 20065	0712-4600 HA 0712-4600 HA 0712-600 HA 0701-0600 HA 0701-0600 HA 0701-0600 HA 0701-0600 HA 0701-0600 HA 0701-0600 HA 0701-0600 HA 0701-0600 HA 0702-0600 HA 0703-0600 HA 0703-0600 HA 0703-0600 HA 0703-0600 HA 0703-0600 HA				
Prevent interfactors		2006051	0T04:05:00 NA 0T04:12:00 NA				
 be skoldkov (5-s woldkov (5-s woldk		Preview A	Annotations				
Message Pilopin Date 1 Facther Engigitet, rices unic adu/2007/am/fdow/data/40320/3b-1aix-48/db bc/8-3b0x8dba2e as /mpl/tyberregreater/1744448080/br edu unic rices optiont 2006-07-06 151:417.105 1 Could not load latest version of context from disk. edu unic rices optiont 2006-07-06 151:417.105 6 K.S. moning message: ContextSinup, Wolf-Wy in edu unic rices optiont regretor unimessages org alignes ragi 2006-07-06 151:018.386 1 Could not load latest version of context from disk. edu unic rices optiont 2006-07-06 151:018.386 1 Could not load latest version of context from disk. edu unic rices optiont 2006-07-06 151:018.386 1 Could not load latest version of context from disk. edu unic rices optiont 2006-07-08 124:255:502 1 Countert dist trapic. obtion transactores optiont grace(size carg) 2006-07:02 13:42:55:502 1 Countert dist trapic. Obtion transactores from disk. edu unic rices optiont edu unic rices optiont 2 Could no load latest version of context from disk. optiontergratoriul messages org alignes cargos normation context dist trapic option context from disk. 6 MAST missing profeferenc	D= Workflow D+ Workflow Graph 🖉	Tool Execution 😢 Error Li	0g 11				JI 6, 16 🗙 🗟 🕼 7 🗝
I Active Hanging Annual Annua	Message				Plug-in	Date	4
Control data tests version of context from dak. educan crass-geterrit 2006-70-015 10:18.896 ALS missing metasage: ContextState, JouRendow J, Valance A, Opermetagratoriu messages organizatione and tests version of context State, Journal 4, Operameters, and 2 blob to new format. educan crass-geterrit 2006-70-015 10:18.396 Locavertal 31 triples, O documents, O state data, O parameters, and 2 blob to new format. educan crass-geterrit 2006-70-015 10:18.396 ALS missing message: ContextState, Journal 4, Operameters, and 2 blob to new format. educan crass-geterrit 2006-70-015 10:18.396 ALS missing message: ContextState, Journal 2006-70:13 42:05 500 Context and the state version of context from dak. educan crass-geterrit 2006-70:01 34:255 500 Context and Targets, Dolonders, Lotol data, Operameters, and 2 blob to new format. educan crass-geterrit 2006-70:01 34:255 500 Context and the professioner. educan crass-geterrit 2006-70:01 33:02 296 Context and the professioner. educan crass-geterrit 2006-70:01 33:02 296 Context and the professioner. fetched the professioner to: non-manifolyhopertegratory-barrierg	I Fetched http://cet.ncsa.uiuc.e	du/2007/workflow/data/40	3203b3-1a3c-4dbd-bc	29-3b0e58d8a2e2 as /tmp/cyberintegrator/1749463809.tr	edu.uiuc.ncsa.cyberint	2008-07-09 15:14:17.105	
A SE manage message: Construint, Volchely in educur.neta cybernetegratorium messages org. (2004-700 31 31:01.8.3.84) Converted 31 Step data, 50 ste	Could not load latest version r	of context from disk.			edu.uuc.ncsa.cyberint	2008-07-09 15:10:18.806	
Countered 28 triples, 6 documents, 0 step data, 0 biol data, 0 parameters, and 2 blobs to new format. edu usc. rcsa. cyberrit: 2006-07-02 13 42:35.592 countered 28 triples, 0 documents, 0 step data, 0 parameters, and 2 blobs to new format. edu usc. rcsa. cyberrit: 2006-07-02 13 42:35.592 countered 28 triples, 0 documents, 0 step data, 0 parameters, and 2 blobs to new format. edu usc. rcsa. cyberrit: 2006-07-02 13 42:35.592 countered 28 triples, 0 documents, 0 step data, 0 parameters, and 2 blobs to new format. edu usc. rcsa. cyberrit: 2006-07-02 13 42:35.549 countered 28 triples, 0 documents, 0 step data, 0 parameters, and 2 blobs to new format. edu usc. rcsa. cyberrit: 2006-07-02 13 32:02 286 countered 28 triples, 0 documents, 0 step data, 0 parameters, and 2 blobs to new format. edu usc. rcsa. cyberrit: 2006-07-02 13 32:02 286 countered 28 triples, 0 documents, 0 step data, 0 parameters, and 2 blobs to new format. edu usc. rcsa. cyberrit: 2006-07-02 13 32:02 286 for chain step preferences: rescue to edu;2007/new/flow/data/4219/046-681-4d0-423-37-02-0586-687 rescue to edu;2007/new/flow/data/42192198-6801-4d0-423-37-02-0586-687 rescue to edu;2007/new/flow/data/421921924-687 rescue to edu;2007/new/flow/data/42192198-6801-4d0-423-37-02-0586-687 rescue to edu;2007/new/flow/data/421921924-687 rescue to edu;2007/new	A NLS missing message: Conte	stSetup_NoDerby in: edu.	uiuc.ncsa.cyberintegra	ator <i>ui m</i> essages	org.eclipse.osgi	2008-07-09 15:10:18.368	
Could not land latest version of context fram disk efforts consistence 2006-07-02 134:25:03-092 Al-ST mission pressess; Control-States, Discher je version cross spherer 2006-07-02 134:25:03-092 Al-ST mission pressess; Control-States, Discher je version cross spherer 2008-07-02 134:25:03-092 efforts assing preferences to: framing between the state of the state of the statest and the statest assistence a	Converted 28 triples, 0 docum	nents, 0 step data, 0 tool	data, 0 parameters, ar	nd 2 blobs to new format.	edu.uiuc.ncsa.cyberint	2008-07-09 15:10:18:345	
14.5 missing message: ContextSetup, NoDMarby in: eduulux nosa optimistagrational messages org. 2006;67:40:23:42:35:330 1 Convented 31 triples, 0 documents, 0 step data, 0 tool data, 0 parameters, and 2 blobs to new format. eduulux nosa optimistagrational messages org. 2006;67:40:23:42:35:330 1 Convented 31 triples, 0 documents, 0 step data, 0 tool data, 0 parameters, and 2 blobs to new format. eduulux nosa optimistagrational messages gradges cores mutitives gradges cores mutitives gradges cores mutitives restored http://cet.nosa.uci.edu/2007/ben/Regulate/2016;680:14006.aptimistagration/sectored at physics/sectore.nutrients; org.ed/sec.exe.uci.edu/2007/ben/Regulate/2016;680:14006.aptimistagration/sectored at physics/sectore.nutrient; org.ed/sec.exe.exe.uci.edu/2007/ben/Regulate/2016;680:14006.aptimistagration/sectored at physics/sectored at	I Could not load latest version of	of context from disk.			edu.uiuc.ncsa.cyberint	2008-07-02 13:42:35.952	
Converted 29 triples, 6 documents, 0 step data, 0 tool data, 0 parameters, and 2 blobs to new format. Problems saving preferences: org.eclipse core nummer; 0 org.eclipse; 0 org	A NLS missing message: Corte	stSetup NoDerby in: edu	uluc nosa cyberintegri	ator ui messages	ora eclase.osai	2008-07-02 13:42:35:520	
Problems saving preferences. ang edpise core rutim 2008-01-02 13:302 295 Topolems to internatively bennegator() bennegator	Converted 28 triples 0 docum	nents, 0 step data, 0 tool	data. 0 parameters a	nd 2 blobs to new format.	edu viuc nosa cyharine	2008-07-02 13:42:35 499	
Exception saving preferences for Animetrallyhybeintegration/cybeintegratio/cybeintegration/cybeintegration/cybeintegration/cybeintegratio	Problems saving preferences				oro eclose core nution	2008-07-02 13:33:02 295	
Fetchel http://etc.ncsa.ucc.ebu/2007/workflow/data/4e219048-6601-4dx6-a233-670e3c6964e7 as //mg/cyberntegrator/antual.tet eduulor.ncsa.cybernt.2008-07-021321352.387 set.ed.nov.	Exception saving preferences	to: hometralphicyberinte	constor/Cyberinternator	rhorkspacel metadatal phonsion eclipse core notime?	orn eclipse equiper on	2008-07-02 13:33-02 288	
Considering (Considering Considering Considering Considering Considering Considering Considering Considering Considering Constraints) Considering Constraints Con	Estrad http://cet.ncva.usic.a	du/2007/workflowid=taite	219048-6801-4454-0	233.670e3r99def7 as #mn/ryberintegrator/mining/	educing ness cohoring	2008-07-02 13:31:52 367	
and the set	A 41 PT	and a set of the set of the set of the set		and a second and the second and second at the second and the secon			
800 121 1008							8M of 10M

the Edit Window Help						
Open Perspective						1.12
Show View 🔹 🏨	Annotations					12
🗟 Data 3	Contaxt	ocument. Wind Speed 12	Annotations			0 4 40
Park Bit Bit and Perspective Bit and Bit and	Dara Mo Steps Teol Secutors Tools University Craph 202 202 202 202 202 202 202 202 202 20	mm313/11 (2000) 1.3 mm313/11 (2000) 1.4 mm313/11 (2000) 1.40 mm313/11 (2000)				
	Pre	wew Annotations				
D= Workflow D= Workflow Graph € Too	Execution 🐑	Error Log 11				2010/06/88 00 07 77 7
Message				Plug-in	Date	
Fetched http://cet.ncsa.ukuc.edu/	2007/workflow/d	lata/403203b3-1a3c-4dbd-bo	c28-3b0e58d8a2e2 as /tmp/cyberintegrato	/1749463909.tr edu.uuc.ncsa.cyberir	2008-07-09 35:14:17.205	
Could not load latest version of c	ontext from dis	k.		edu uuc ncsa cyberir	2008-07-06 15:10:18 806	
A NLS missing message: CortextS	etup_NoDerby i	n: edu uiuc nosa cyberintegr	ratorui.messages	org.eclpse.org	2008-07-09 15:10:18:368	
1 Converted 28 triples, 0 document	ts. 0 step data.	0 tool data. 0 parameters, a	and 2 blobs to new format.	edu siuc nosa cyberir	2008-07-09 15:10:18:345	
Could not load latest version of c	ortext from dis	k,		edu uluc nota cyberin	2008-07-02 13:42:35.852	
A NLS missing message: CortextS	etup_NicDerby i	n edu.uuc.ncsa.cyberintegr	ratorulmessages	org.eclpse.osgi	2008-07-02 13:42:95:520	
I Converted 28 triples. 0 document	es. O step data.	O tool data, O parameters, a	and 2 blobs to new format.	edu uiuc nosa cyberin	1 2006-07-02 13:42:35.499	
Problems saving preferences.				org.eclpse.core.runti	7 2008-07-02 13:33:02.295	
Securition saving preferences to:	hometralphycy	berintegrator/Cyberintegrato	or/workspace/.metadata/.plugns/org.eclips/	core.runtime/1 org.eclpse.equinox.p	e 2008-07-02 13:33:02.288	
Fetched http://cet.ncsa.u/uc.edu/	2007/workflow/d	ata/4e219048-6801-4d0d-a	233-670e3c80def7 as /tmp/cyberintegrator	retval.txt edu.uur.ncsa.cyberir	e 2008-07-02 13/31/52/367	
	- 1 C					
						5H of 10H

Figure 2. Accessing perspectives (top) and views (bottom) via the main menu->window options.



Figure 3. Open perspective by selecting one of the saved perspectives in the dialog launched by "main->window->open perspective->other".

* •		Cyberintegrator			8
Eile Edit Window Help					
			E C		**
🗟 Data 🕅 🔍 🗸	Document: Wind Direction 🗱	Annotations		û v e	
type filter text	20060510T12:00:00 4.000 20060510T12:06:00 NA 20060510T12:2:00 A 500				-
Wind Direction (DONE)	20060510T12:18:00 NA				
Wind Speed [DONE]	20060510T12:24:00 NA 20060510T12:30:00 4.400				
🗟 retval (DONE)	2006453 10112:3:4:500 MA 2006453 10112:3:4:500 MA 2006453 10112:3:4:500 MA 2006453 10112:3:4:500 MA 2006453 101113:4:500 MA 2006453 101113:2:00 MA 2006453 101113:2:00 MA 2006453 101113:2:00 MA 2006453 101113:4:00 MA 2006453 10113:4:00 MA 2006453 10112:3:00 MA 2006453 10112:3:00 MA				
Tools 22	20060510102:30:00 4.200 20060510T02:36:00 NA				
tune filter text	20060510T02:42:00 3.700 20060510T02:48:00 NA				
27 TCOON data (v1) 27 Transform TCOON wind data (v1)	2006051070254:00 4.200 20060510703:05:00 4.200 20060510703:05:00 NA 20060510703:12:00 3.300 20060510703:24:00 NA 20060510703:24:00 NA 20060510703:36:00 NA 20060510703:36:00 NA				
	20060510T03:54:00 4.700 20060510T04:00:00 4.700				
	20060510T04:06:00 NA 20060510T04:12:00 NA				
	Preview Annotations				1
🕽 - Workflow 🕄 🗦 Workflow Graph 🦉 Too	Execution 😲 Error Log			-	
General Information		Descrip	otion		1
Name: Workflow 1				*	
Creator: Anonymous					
Date: jun 10, 2008 4:08:41 PM					
Contributors:				-	
	Reset Apply	4		•	
			5/	N 01 10M	ø

Figure 4. An example of a perspective with the list of data sets (left upper view), the list of tools (left lower view), visualization of the content of the Wind Direction data set (right view) and the annotation panel for describing the workflow and creating metadata about the workflow (bottom view).

🗿 Cyberintegrator			
File Edit Window Help			
] Login			🗈 🈭 CyberIntegr
🚘 Tools 励 Data 🕱	~	🛈 Workflow : SWMMORGN.INP 🛛 🐌 Annotations	
type filter text		Misurdi	zation
		52380	Visualization
testdiams.txt [DONE]		Visualization	text/plain W/Visualtzes
MARK_MM.DAT 879 bytes [IMPORTED]		Visualization	rect:/workflow/visuali:
🕀 🕞 SWMMORGN.INP 52380 bytes [IMPORTED]		Cet:Document	cet:/tags/1c7460609ad4b8522e
- DONE]		Visualizati	Dauzes~dc;creator on wt/%kep;kingout/Valuelizes
SWMMPROJ.OUT [DONE]		Visualization	Anonymous
SWMMPROJ.RPT [DONE]		SWMMORGN INP	Vieualization
best_pop.out [RUN] for the to have [Mapon TED]			(Suchzadori)
alphaleaarch in EE7 button [IMPORTED]			
globalsearch.in 567 bytes [IMPORTED]	-	Proview Attributes Tupelo View Appotations	
		The view Actinuates Topelo view Annotations	
🐉 Workflow 🗊 Workflow Graph 🛛 🦉 Step Creator 🥺 Error Log			
test.pdf) ► <mark>後 acrobat</mark>		📔 output pdf) 🛛 🕞 🚱 acrobat)	
surtreshold.txt)			
super URL web page			
fraction.txt			
integration in the state in the			
		pipesdown.txt	<u>G</u> .
SWMMORGN.INP)			
dir output			
			•
			14M of 18M 🔟

Figure 5. An example of a perspective with the list of data sets (left view), the visualization of metadata about the data set called SWMMORGN.INP (right view) and the visualization (a workflow graph) of all activities (bottom view).

計(Cyberintegrator	_ 🗆 ×
Eile	a Edit Window Help	
Log	igin 🖉	CyberIntegr
-	🛈 Workflow : paretosol 🛛 🖉 🎩 Annotations	
	inivars.btt SWMMPROJ.INP ORGDIAM.btt MARK_MM.DAT best_pop.out	
5		
	set;/workfakk/maik/maik/maik/maik/maik/maik/maik/m	
	naretron)cet:/workflow/tool/r	asīnput ►(G.pin)
	Anonymous detereation	
	ret:/workflow/Tool	
	dc:description set://workflow/isAssociated	
	(cet:/tags/4a8af6219f943055fb8a0469f2d5772b1b29de7d51) Oct 23, 2007 8:35:10 AM	
	(solutions generator) (demo workflow)	
	Help Parameters Attributes Tupelo View Annotations	
	ခြာမ Workflow ခြာမ Workflow Graph ဆိ 🖉 Step Creator တြဲ Error Log	- 0)
	🖾 text off	
		-6
	i output URL web page	
	fraction.txt	
	🥐 netstat	~
	SWAMACT SWAMACT SWAMACT SWAMACT SWAMACT	
	kplorer F	
	•	
		21M of 24M 📋

Figure 6. An example of a perspective with visualizations of metadata about the tool called paretosol (top view) and a workflow graph of all activities (bottom view).

1.3 Preferences

Properties of Cyberintegrator can be accessed and modified by selecting main->windowpreferences. There are properties related to

- editable metadata entries in attribute editor
- authentication information for logins
- additional views visualizing metadata and attributes
- information for connecting to a Liferay-based portal system at NCSA
- log settings during Cyberintegrator executions
- settings for browsing context
- local and remote storage of context
- visualization of user actions in the view displaying a workflow graph.

These properties are explained in Figure 7 through Figure 13.

* •	Preferences	$ \uparrow \Box \times$
type filter text	Authentication	⇔ ∙ ⇔.
 Authentication General NCSA Logging Preferen Remote Execution Tupelo Preferences 	Authentication Preference Page Service Provider	
•	Restore De	faults <u>A</u> pply
	ок	Cancel

Figure 7. Preference page related to the form of authentication. If "Context Login" would be selected then a user can register and authenticate later using his/her user identification (ID) and password when working with Cyberintegrator. The user ID replaces the anonymous default creator and is used when the work in Cyberintegrator is performed (e.g., creation of workflows, data sets, tools or annotations). If "Liferay" is selected then the authentication is performed against the user ID and password in the NCSA CyberCollaboratory.⁴

4

 $http://cleaner.ncsa.uiuc.edu/cybercollab/c/portal/layout?p_l_id=1.1$

	Preferences	s A D X
type filter text	Liferay	
 ✓ Authentication Liferay Ø General NCSA Logging Preferences Remote Execution Tupelo Preferences 	Liferay Authentication Liferay Endpoint: Liferay User ID (key): Liferay Password: Liferay Company ID	Preference Page http://nyssa.ncsa.uiuc.edu:10010 15407 ********* 100912
4	Þ	Restore Defaults Apply OK Cancel

Figure 8. Preference page containing authentication information to connect to a Liferaybased portal system at NCSA called Cybercollaboratory.⁵

5

•

 $http://cleaner.ncsa.uiuc.edu/cybercollab/c/portal/layout?p_l_id=1.1$

* •	Preferences	$ \bullet \square \times$
type filter text	General	⇔ ∙ ⇔∽
 ✓ Authentication Liferay ✓ General Attribute Editor ♦ NCSA Logging Preferen Remote Execution ♦ Tupelo Preferences 	Workflow Graph Show user actions? Info Viewer Use preview if available? Show attributes? Show tupelo view? Tools Show all versions in context view? Show only latest version in tools view? Temp Space Temp Folder: /tmp/cyberintegrator	
4	Restore <u>D</u> efaults	<u>A</u> pply
	ОК	Cancel

Figure 9. Preference page that allows (a) hiding or showing (check box labeled as "Workflow Graph" is checked by default) user actions, such as import data/tools or preview data, in the view displaying a workflow graph, (b) using previews of large files (the "Use preview" checkbox) and (c) adding two extra views to the collection of all available views. The Tupelo view is shown in Figure 5.

☆ ◎	Preferences	↑ □ ×
type filter text	Attribute Editor	⇔ ▼ ⇔∨
Authentication	List of predicates that can be modified	
▽ General	http://purl.org/dc/elements/1.1/title	New
Attribute Editor	http://purl.org/dc/elements/1.1/description	
NCSA Logging Preferen		Eemove
Remote Execution		<u>U</u> р
v lupelo Preferences		Down
4	Restore <u>D</u> efaults	Apply
	ОК	Cancel

Figure 10. Preference page related to editable metadata entries. The "Show Attribute" checkbox refers to the Attribute View and the visibility of attributes. The list of entries under the label "list of predicates that cane be modified" refers to those predicates in the RDF triples (subject-predicate-object) that are enabled for editing. We do not recommend adding additional editable predicates!

	Preferences	
type filter text	NCSA Logging Preference Page	⇔ ∨ ⇔∨
 Authentication General Attribute Editor NCSA Logging Preferen Console Logging Pref∉ Eclipse Logging Pref∉ Log4J Logging Prefer 	Global Log Level ALL	
Popup Logging Prefe Remote Execution ♪ Tupelo Preferences		
4		Restore Defaults Apply OK Cancel

Figure 11. Preference page for selecting different levels of logging, along with the global log level. Within this subtree, you can select different types of logging, such as Console logging, Eclipse logging, Log4J logging, and Popup logging. Each has its own preferences that you can customize to your liking.

🛊 Preferences		IX
type filter text	Tupelo Browser 🔶 👻	ð 🔻
Authentication Liferay General Attribute Editor Log Provider Tupelo Preferences Tupelo Context	Context Preference Page Layout : SpringLayoutAlgorithm Depth : i <lii< li=""> i <lii< li=""></lii<></lii<>	-
	OK Cancel	

Figure 12. Preference page related to browsing the context (metadata). The settings define the appearance of the graph visualizing the metadata retrieved from Tupelo. Layout – automatic layout of the graph. Depth – how many layers deep the sub-nodes should be shown. Direction – filters graph showing only sub-nodes that contain selected directions Subject->Predicate->Object of the predicates.



Figure 13. Preference page with the four tabs related to local and remote storage of metadata. The current setting in the top left screen is for local context. The location of the local context is defined in the second tab (top right). In the first tab (top left), the checkboxes "WebDAV/Sesame Context" and "mySQL Context" should be checked when the end product of the work in Cyberintegrator should be shared with a community. In the current prototype, Sesame is the computing infrastructure supporting secure interactions among the client (Cyberintegrator user), NCSA server, and NCSA database tiers. WebDAV is the Web-based Distributed Authoring and Versioning software managing metadata information on remote servers.

Chapter 2 - Getting Started

2.1 New Workflow and Login

After starting Cyberintegrator, one can start with creating a new workflow and logging into the system if needed.

(1) Starting a new workflow

Select File->New->Workflow from the main menu. A new dialog will appear as shown in Figure 14. The title (name) of the workflow will be associated with all activities in Cyberintegrator. Figure 15 shows the appearance of Cyberintegrator after creating a new workflow.

		New Workflow	_	↑ □ X
New Work	flow			
Title				
			risish (Connel
			Einish	Cancel

Figure 14. A dialog invoked by File->New->Workflow that allows starting a new workflow.

* •	Cyberinteg	rator		• - 6 ×
Eile Edit Window Help				
in n			1	30
🖬 Data 🖾 🗸 👻 🗖	I Info: N/A 😫 👘 Annotations			∂ ⊽ ■ □
type filter text	1			
Sools Si Cols				
worknow 25 Con Worknow Graph @ 1001 Ex	ecution Vierror Log			0
General Information		Description		
Name: New Workflow				E I
Creator: Anonymous				
Contributors: Jul 9, 2008 3:44:07 PM				
	Reset Apply			
		1		
				of 10M

Figure 15. Cyberintegrator appearance after creating a new workflow.

(2) Logging as a known user.

When Cyberintegrator is started, the user name is automatically set to anonymous. If a user would like to use his/her known login then the login can take place in two ways. The first way is based on an already existing user login in the NCSA Liferay based portal system called Cybercollaboratory. The second one is based on a local context where a user account is created specifically for Cyberintegrator login. The Liferay or local context type of login is defined in the authentication preference dialogs, accessible by going to Window >> Preferences in the main Cyberintegrator window.

Depending on the authentication preference dialog, a login dialog is invoked by clicking on the "Login" button in the top left corner of Cyberintegrator right below the File menu. One can then login by authenticating against the Liferay based portal (see Figure 16 (top) or against the local context (see Figure 16 (bottom).

* •	Liferay Login	↑ ×
Please enter your Lif	eray account information to auth	enticate.
Login:		
<u>P</u> assword:		
	Login Ca	ncel
* •	Context Login	T X
Email address is requir	ed.	
Email address:		
Password:		
Following field is only re	equired if registering a new account.	
Fullname:		
	OK Cancel Regis	ster

Figure 16. Based on the settings for authentication shown in

Figure 7, a user login is enabled either via Liferay (top) or the local context (bottom). The appropriate dialog is invoked by clicking on the "Login" button in the top left corner of Cyberintegrator right below the File menu. A new Liferay account can be created at the NCSA CyberCollaboratory web page. A new local context account is created by pressing the button "Register" (bottom figure). A user will see his/her user ID by viewing the Workflow View (previously anonymous).

2.2 Importing and Exporting Data, Existing Tools and Workflows

Cyberintegrator enables users to easily import and export existing data, workflows and tools into and from a software/data unifying environment.

(1) Importing Data:

Importing data can be performed by either dragging followed by dropping the data files into the data pane (see Figure 18) or by selecting File->Import followed by the choices shown in Figure 17. In both cases, the mime⁶ type of data can be determined automatically based on the

⁶ From Wikipedia: "**Multipurpose Internet Mail Extensions** (**MIME**) is an Internet Standard that extends the format of <u>e-mail</u> to support:

suffix of the data files. For example, JPG or TIF suffix will indicate image mime types while TXT or CSV would indicate plain/text types. Application/octet-stream should be used for application (binary) data types. When a user chooses the drag/drop option for importing data files then the data mime type will be determined automatically by default and if the mime type could not be determined then a dialog will be shown. If a data set contains metadata about the type, select the "Data including metadata" option. Otherwise, you can select the "Data no metadata" option and specify the type yourself.

* •	li	mport	_	↑ □ ×
Select				
Select an import so	UFCO:			
<u>select an import so</u>	urce.			
type niter text				
🗟 Data includin	g metadata			
🗟 Data no meta	adata			
🐎 Workflow				
🚘 XML Tool(s)				
	< <u>B</u> ack	<u>N</u> ext >	Einish	Cancel

Figure 17. The dialog for importing data, workflows or tools.

If a data file imported is larger than a pre-set file size threshold then a preview will be automatically created. Figure 18 illustrates this example when the selected image is imported and a preview can be viewed immediately.

- text in <u>character sets</u> other than <u>US-ASCII</u>;
- non-text attachments;
- multi-part message bodies; and
- header information in non-ASCII character sets. ". All media mime types can be found at http://www.iana.org/assignments/media-types/.

🋊 Cyberintegrator		_ _ _ _ ×
File Edit Window Help		
Login		🗈 😭 CyberIntegr
🖬 Data 🛛 🗖 Steps 🛛 🖓 🗖	🛈 Document: Preview Image 🛛 順 Annotations	
type filter text		
⊟ tiff ⊢ ILUTMsubSamp30.tif [IMPORTED]		
Preview Image		
	State States	
type filter text		
	Preview Annotations Tupelo View Attributes	
🕽 Workflow 🗊 Workflow Graph 😫 🙋 Step Creator		- 0
(Imported)	► 🔜 ILUTMsubSamp30.tif)	
Review Image	> 🔂 Visualization	
		33M of 37M

Figure 18. Example of importing data. Importing a land use/land cover map to the data pane, selecting the preview image for viewing in the right pane and viewing the workflow graph of activities in the bottom pane.

(2) Exporting Data:

The export of any data file is enabled (a) by File->Export menu (see Figure 19), (b) by right mouse click on the data file and selecting the export option (Figure 20) or (c) by dragging and dropping the file to a File explorer.

🋊 Cyberintegrator	
File Edit Window Help	
Login	😰 🈭 CyberIntegr
🖬 Data 🛛 🗖 Steps	🛈 Document: ILUTMsubSamp30.tif 🖄 🐠 Annotations 🔒 🍸 🗖 🗖
type filter text image image	
	Preview Annotations Tupelo View Attributes
🕽 - Workflow 🕞 - Workflow Graph 😣 🧵 Step Creator	
Imported	> ILUTMsubSamp30.tif) > ILUTMsubSamp30.tif)
Preview Image)	Visualization
	33M of 37M
🎝 Start 🗀 2 Wind 🗸 🍠 4 Inter 🗸 🤿 en	viroci 🗹 2 Micro 🌸 Cyberin 🦳 99% 🗗 🕊 🔍 📲 🖓 🕄 🖓 🕅 🖳 👫 8:26 PM

Figure 19. Example of exporting data using the right mouse click on the data file.

🛊 Export	_ 🗆 🗙	🛊 Data Export Wizard 📃	
Select	2	Select data to export.	-
Select an export destination:		Show all data in context?	
type filter text		ILUTMsubSamp30.tif Runner.jpg	
- 📅 Data - 🚘 Tool			
1		Directory: C:\PeterB\Projects\TestData\FileFormats	
< Back Next > Finish	Cancel	< Back Next > Finish Cance	

Figure 20. Example of exporting data using the File->Export option.

(3) Importing Existing Tools:

Tools are viewed as a "black box" with inputs, outputs and parameters. This model is used for creating xml descriptions of tools. Two examples of xml descriptions of tools are provided in Figure 21.



Figure 21. Two examples of XML files that represent a Notepad tool (top) and SWMM tool (bottom). The Notepad tool take original text, opens the text in Notepad and then saves out the modified text. The SWMM tool encapsulates the Storm Water Management Model (SWMM), its inputs and outputs.

In order to create tools that could be executed in Cyberintegrator, one has to create the xml description. We will explain later how these xml descriptions can be created using wizards.

Assuming that there are exist tools (xml files describing tools); one can import the tools by either dragging followed by dropping the xml files into the tool pane (see Figure 23) or by selecting File->Import followed by the choices shown in Figure 22



Figure 22. Example of importing tools by invoking dialogs via File->Import.



Figure 23. Example of importing the tool called Notepad.

(4) Exporting Tools:

The export of any tool is enabled (a) by File->Export menu (see Figure 24), (b) by right mouse click on the tool and selecting the export option or (c) by dragging and dropping the file to a File explorer.

🛊 Export	<u> </u>	Export tool(s)	
Select		Select Tool(s) to Export	
Select an export destination:		☑ Show all tools in context	
type filter text		echo	-
		Internet explorer Matrix inverse	
Stew Workflow		SWMMC1 paretosol	
		test NSGAx	
		echo	
		Time Series Plot	
		2D Plot	
		Minipie test matrix MS Word	
		Internet Explorer netstat	•
		Filename: C:\Documents and Settings\pbajcsy\Desktop\notepad.xml	
< Back Next > Finish	Cancel	< Back Next > Finish Cance	!

Figure 24. Example of exporting tools by invoking dialogs via File->Export.

(5) Importing and Exporting Existing Workflows:

Importing and exporting workflows is described in Figure 25 and Figure 26.

🛊 Import	<u> </u>	Norkflow Import Wizard	_ 🗆 🗙
Select	Ľ	Select workflow file to import.	
Select an import source: type filter text		Filename: C:\Documents and Settings\pbajcey\Desktop\workflowTest.zip	
< Back Next > Einish	Cancel	< Back Next > Finish C	ancel

Figure 25. Example of exporting workflow by invoking dialogs via File->Export.



Figure 26. Example of exporting workflow by invoking dialogs via File->Export.

2.3 Viewing and Annotating Data, Tools and Workflows

Cyberintegrator comes with a set of functionality for visualizing text/ASCII and image data, as well as metadata and annotations about data. In addition, data, tools and workflows can be annotated with unstructured text or text following the RDF metadata format.

(1) Viewing Image and Text/Plain Data:

Once a data set has been imported, it can be viewed by right mouse clicking on the data in the data pane and selecting "Open in New Window" (see Figure 27). Multiple viewers and viewing options can be accessed from the right upper corner of the window. Figure 28 shows the options for image data and Figure 29 shows the options for text/plain data.

🛊 Cyberintegrator	_	8 ×
File Edit Window Help		
Login	🖽 🕞 CyberIntegr	
🔂 Data 🛛 🗖 Steps 💎 🗖 🗖	🕼 Annotations 🛈 Tool: notepad 🛈 Document: params_pop.out 🛈 Document: ILUTMsubSamp30.tif 🖄	
type filter text		
text final pop.out [IMPORTED] Freview Text Freview Text		
	Preview Annotations Tunelo View Attributes	
1 9 a Workflow 22 9 a Workflow Croph 2 Ctop Cropter		-
Name: new workflow	Description	
Creator: Anonymous		
Date: Nov 1, 2007 8:18:25 PM		
Contributor:	•	
Booot Apply Vieweline		-
Reset Appry Visualize		
	5M of 37M	Û
🏄 Start 🔁 2 Wi 🚽 🥭 4 Int 🚽 🛢 enviro	🗹 2 Mic 🔹 🖓 Cyber 🦉 error W 🦳 🤧 🖓 🖬 🕄 🎔 🕅 😤 🎦 9:48	РМ

Figure 27. A data file can be viewed by right mouse clicking and selecting the option "Open in New Window".

🋊 Cyberintegrator		_ 8 ×
File Edit Window Help		
Login	E 🔐	CyberIntegr
🖬 Data 🛛 🗖 Steps 🛛 🖓 🗖 🕼	Annotations 🛈 Tool: notepad 🛈 Document: Runner.jpg 🛛	
Itype filter text image ipeg image image	Vie Preview Image	ta
Pre	view Annotations Tupelo View Attributes	
🕽 •• Workflow 🛛 📚 •• Workflow Graph 🖉 Step Creator 🗖	Tupelo Browser	
General Information	Description	
Name: new workflow Creator: Anonymous Date: Nov 1, 2007 8:18:25 PM		*
Contributor:	×	
Reset Apply Visualize	I	V V
		27M of 37M 🛛 📋
🏄 Start 🗀 2 Wi 🖌 🍠 4 Int 🖌 🏓 enviro	📝 2 Mic 🔻 🌸 Cyber 🦉 errorW 🧾 99% 🔰 🕿 🔍 🖼 🖏 🎔 🕅 景	. 🎦 9:43 PM

Figure 28. An image data viewer can show either the original image or the preview image. The setting for image viewers can be accessed by left mouse clicking on the triangular shape pointing downward in the right upper corner of the visualization window as illustrated here.

🋊 Cyberintegrator		_ _ _ _ _ _
File Edit Window Help		
Login		🗈 🏫 CyberIntegr
🖬 Data 🛛 🗖 Steps 🛛 🖓 🗖	Annotations (1) Tool: notepad (1) Document: Runner.jpg (1) Docum	nent: params_pop.out 🛛 🔒 🌱 🗖
type filter text Itor Itor	# This file contains information about inputs as read by the program Population size = 4000 Number of generations = 20 Number of constraints = 0 Number of constraints = 0 Number of real variables = 74 Lower limit of real variable 1 = 1.00000000000000E+001 Upper limit of real variable 2 = 1.0000000000000E+001 Upper limit of real variable 3 = 1.1000000000000E+001 Upper limit of real variable 3 = 1.200000000000E+001 Upper limit of real variable 3 = 1.20000000000E+001 Upper limit of real variable 3 = 1.20000000000E+001 Upper limit of real variable 4 = 1.20000000000E+001 Upper limit of real variable 4 = 1.20000000000E+001 Upper limit of real variable 5 = 1.40000000000E+001 Upper limit of real variable 6 = 1.40000000000E+001 Upper limit of real variable 6 = 1.40000000000E+001 Upper limit of real variable 7 = 1.40000000000E+001 Upper limit of real variable 7 = 1.50000000000E+001 Upper limit of real variable 7 = 1.500000000000E+001 Upper limit of real variable 7 = 1.5000000000000E+001 Upper	• TextView Viewer ▶ HTMLView Data ▶
The Workflow 🕅 💭 Workflow Graph 🖉 Step Creator	Tupelo Browser	
General Information Name: new workflow Creator: Anonymous Date: Nov 1, 2007 8:18:25 PM	Description	×
Contributor:	~	
Reset Apply Visualize	1	
		33M of 37M
🏄 Start 🗀 2 Wi 🚽 🥭 4 Int 🖌 🛢 enviro.	📝 2 Mic 🗸 🌸 Cyber 🦉 errorW 🧕 99% 🕴 🏶	🛚 < 🔊 🖬 🕄 🎔 M 🖳 🎦 9:47 PM

Figure 29. A text/plain data viewer can show either the original text or the preview text (top lines only). The setting for text/plain viewers can be accessed by left mouse clicking on the triangular shape pointing downward in the right upper corner of the visualization window as illustrated here.

(2) Viewing Metadata:

There are several views of metadata that are either enabled by the default perspective or can be enabled by selecting from the main menu Window->Show View. One can see the textual or graph views in Figure 30 through Figure 33.

* Cyberintegrator			_ = ×
File Eait Window Help		Tri Si Ocharin	toor
			eyr
Data X Steps	Annotations (1) Tool: notepad (1) Docu	ument: cat.gif 🛛	
type filter text	Predicate 💌	Object	-
E-C image	tag:tupeloproject.org,2006:/2.0/files		
gif 🚺	canRead	true	
🗄 👼 cat.gif [IMPORTED]	canWrite	true	
🖬 Preview Image	hasName	5a0510ffab9bc5ec1cdfff8e00627d923e02eb4e	
🚊 💼 jpeg	isDirectory	false	
E 🔂 Runner.jpg [IMPORTED]	isFile	true	
📕 🚊 👘 tiff 🛛 🛁	isHidden	false	
⊞🔂 ILUTMsubSamp30.tif [IMPORTED]	lastModified	2007-11-02T03:10:20Z	
e 🔁 text	length	12477	
📄 🗁 plain 🔤	📄 🗉 tag:cet.ncsa.uiuc.edu,2007:/workflow,	/da	
📄 🖹 params pop.out [IMPORTED]	preview	tag:cet.ncsa.uiuc.edu,2007:/workflow/data/e13f69a4a0c8ce	:3a
🖻 Tools 🛛 🦉 🗸 🗖	size	12477	
type filter text	□ http://purl.org/dc/elements/1.1		
	creator	tag:cet.ncsa.uiuc.edu,2007:/people/anonymous	
	date	2007-11-02T03:10:20Z	
	format	image/git	
	title	cat.gif	
	Add Delete		
	Provine Appointations Tunala View Attribu	tee	
ļ	Preview Annotations Tupero view Attribu		
🕽 🗠 Workflow 🛛 🔁 🗫 Workflow Graph 🖉 Step Creator	🗖 Tupelo Browser		
General Information	Description		
Name: new workflow			*
Creator: Anonymous			
Date: Nov 1, 2007 8:18:25 PM			
Contributor:			
Reset Apply Visualize			_
	4		>
		28M of	37M 🕅

Figure 30. An attribute view (right window) of the metadata about an imported image cat.gif.



Figure 31. A graph view called Tupelo View (right window) of the metadata about an imported image cat.gif.

Figure 32. A graph view called Tupelo browser (bottom window) of the metadata about all workflow activities.

2 O	_	Cyberintegrator		+ - 6 ×
Eile Edit Window Help				
				📑 🤔 Cyberinte
🖬 Data 🕱 🔍 🔍 🗖	🕕 Tool: notepad 🕴 約	Innotations		🔒 🔻 🗖 🖬
type filter text	notepad			
B Output text [DONE]	· General Information			
Uofi-NCSA.jpg [IMPORTED]	Text editor for the Windows	s. Documents edited need to	be saved with the same name (Ctrl-S), not saved as another document. Thi	s will only work under Windows,
petstat output (DONE) petstat output (DONE)	Parameters			
sphestac output [Doive]	✓ Inputs			
	Name	Format	Description	
	input text	text/plain	Text file that needs to be edited	
	- Outputs			
	Name	Format	Description	
	output text	text/plain	Text file after editing.	
≩रे notepad (v1)				
	Info Annotations			
🕽+ Workflow 🗊+ Workflow Graph 💿 Tool Execution 🕴 🗖 Tupelo Browser				- 0
General Information		• Parameter	rs	- -
This section describes general information about the step that will be created.		* Inputs		
Step name: notepad		This section d	escribes inputs used by the step when being executed.	
Tool name: notepad		input text ne	tstat output	v
Fixecution Location		- Outputs		
Execute Step		This section d	escribes outputs generated by the step when being executed.	
			angkan sama	
				13M of 14M 🃋

Figure 33. An information view (right window) of the inputs, outputs, parameters and any other information (metadata) about the tool called Notepad. Very similar information is provided in the bottom window in the Step Creator view. The Step Creator view is for executing the tool/step.

(3) Adding Annotations:

Annotations can be added to data, tools or workflows. Furthermore, the workflow description can be also added as shown and explained in Figure 34.

2 O	cy	berintegrator		* - 8 ×
Ele Edit Window Help				
				😭 🦓 Cyberinte
Data 22	🗖 🗖 🕕 Document: netstat output	Annotations 83		- 0
type filter text	the DEMO	From : Anonymour		
Output text [DONE] Uof-NCSA.jpg [IMPORTED] Antstat output [DONE] netstat output [DONE]	P ⊆ STEPS P ⊆ TOOLS P ⊆ Pool P	Touri, Analysious Subject: Good Tool Date : Jul 10, 2008 2:13:36 PM This tool opens up Gedit and allows you to enter in text.	1)e.
❀ Tools II Evpe filter text Øf gedk (v1) Øf nottepad (v1)				
		•		•
	New Annotation		Reply Submit	
D= Workflow D= Workflow Graph 20 @ Tool Execution 🗖 Tupelo Browse	r			• 0
50 patetat	tat output			
or netstat	tot output			
i internet i	tat output	⇒ 🔐 gedit	> 🔛 Output text	
				13M of 14M 🍵

Figure 34. A new annotation is added by selecting a data set, a tool or a workflow and clicking the button "New Annotations". When the text is completed the button "Submit" should be clicked.

(4) Sorting Data and Tools:

Many times the number of data sets and tools might be very large and there is a need to search for the data and tools of interest. This can be accomplished by filtering data based on tags, names or types as illustrated in Figure 35.

\$ °	Cyberintegrator	• • • • ×
Ele Edit Window Help		
		😭 🎲 Cyberinte
🛱 Data 🕄	🗢 🗖 🕕 Document: netstat output 🕸 🔪 Annotations	0
	Group Dy Optical Advances Annum 4() Local Address State trop Name 0.0.024600 0.0.0.04 LISTEN trop Step 0.0.03901 0.0.0.04 LISTEN trop Total 0.0.03901 0.0.0.04 LISTEN trop Total 0.0.03901 0.0.0.04 LISTEN trop Total 0.0.004 LISTEN trop Total 0.0.004 LISTEN trop 0.0.0122121 0.0.0.04 LISTEN trop 0.0.0122121 0.0.0.04 LISTEN trop 0.0.0122121 0.0.0.04 LISTEN trop 0.0122108.073810 7.2142231805 CLOSE WAT trop 0.192.108.0738010 7.2142231805 CLOSE WAT trop 0.192.108.0738010 7.214223187.9900 CLOSE WAT trop 0.192.108.0748012 K010.1102748 CLOSE WAT trop 0.192.108.0748012 K010.1102748 CLOSE WAT trop 0.192.108.0748012 K11311	
♣ Tools 11 type filter text # gedit (v1) # netstat (v1) # notepad (v1)	Udp 0 0.0.0.0.5353 0.0.0.0.7 C 0 0.0.0.0.5353 0.0.0.0.7 Active LBRX dynamic diskt (several established) unix 2 1.4CC Stream diskt (several established) unix 2 1.4CC StREAM LISTERMO 1360135 Marphotour.8:Skjav@pi unix 2 1.4CC StREAM LISTERMO 1377225 Marphotour.8:Skjav@pi unix 2 1.4CC StREAM LISTERMO 1377225 Marphotour.8:Skjav@pi unix 2 1.4CC StREAM LISTERMO 1358004 Armpicht_photour.8:Skjav@pi unix 2 1.4CC StREAM LISTERMO 1358004 Armpicht_photour.2:Skjav@pi unix 2 1.4CC StREAM LISTERMO 1358004 Armpicht_photour.2:Skjav@pi unix 2 1.4CC StREAM LISTERMO 1358005 Marpicht_photour.2:Ske-0-301502455 unix 2 1.4CC StREAM LISTERMO 1358005 Marpicht_photour.2:Ske-0-302607411724 unix 2 1.4CC StREAM LISTERMO 134213 Marpicht_photour.2:Ske-0-30280504111724 unix 2 1.4CC StREAM LISTERMO	-
🖫 Workflow 🕞 Workflow Graph 🔢 🖉 Tool Exe	ecution 🗖 Tupelo Browser	- 0
international and a second sec	Constant except	rput text
Contractory	Citatus Conson, "GT yeah, " and	STATE A DAY
		13M of 14M 🌐

Figure 35. Sorting/Filtering of data sets can be accomplished by right mouse clicking on the triangular icon in the upper right corner of the Data view and selecting one of the options. The same filtering can be applied in the Tools view.

In order to assign tags to data or tools, one has to select a data/tool and press ALT-T.

	TAG : netstat output	◆ ×
Tag(s):		
	OK Cance	

Figure 36. The dialog invoked after selecting a data set or a tool and pressing ALT-T. The tag will be used for sorting the files.

Note: It is advised to start a new workflow after very many analyses have been completed otherwise rendering of a large collection of activities might be time consuming. See the complexity of the graph rendered in Figure 37.



Figure 37. Visualization of a very large set of analyses in one workflow set that might take significant computational resources for rendering. It is advised to create a set of independent workflows to avoid slow rendering.

2.4 Creating a Tool

There are several types of desktop applications, remote functionalities accessible via web services or library codes written in multiple languages that might become tools. In Cyberintegrator, they are all defined by their inputs, outputs, parameters and executables with the appropriate libraries. The challenge of creating a tool in Cyberintegrator is to simplify the tool creation for a novice.

In order to make the tool creation in Cyberintegrator easer for users, Cyberintegrator provides multiple wizards for tool creation. The wizards are accessible from the main menu File->New as illustrated in Figure 38.

A 9			Cyberintegrator		
Eile Edit Window Help					
New • Cyberintegrator tool					TT Cyberinte
Open Warkflow Command Line Tool					and first excerning
Pename F2 Java Tool Befresh E5 Matlah Tool	~ ~ ~ 1	Tool: notepad 13	Annotations		(A) > + C
Login D= WorkFlow		notepad			
Convert Context		· General Informati	on		
import		Text editor for the Wind	lows. Documents edited need to	be saved with the same name (Ctrl-S), not saved as another docume	nt. This will only work under Windows.
📫 Export		A Parameters			
Exit		- Inputs			
		+ inputs			
		Name	Format	Description	
		input text	text/plain	Text file that needs to be edited	
		· Outputs			
		Name	Format	Description	
		output text	text/plain	Text file after editing.	
 novelan (x1) 		1			
		Info Annotations			
D= Workflow D= Workflow Graph 🦉 Tool Execution 🕫	Tupelo Browser				- c
General Information	and a second		+ Paramete	rs	
This section describes general information about the	he step that will be create	d.	* Inputs		
Step name: notepad			This section d	escribes inputs used by the step when being executed.	
Tool name: notepad			input text ne	tstat output	(v)
Execution Location			a Outente		
Execute Step			This section d	escribes outputs generated by the step when being executed.	
			output text	output text	
					13M of 14M

Figure 38. Tool creation using wizards can be invoked from File->New as illustrated above.

Cyberintegrator version 2 has a rather expansive array of wizards to assist the end user in tool creation. The Command Line Tool wizard is for creating tools based on executable files that could have been run in a command line or a shell. The Java tool wizard is meant for creating tools by interfacing application programming interface of methods written in Java and available in jar files. Finally, a Matlab tool wizard is built for creating tools that are actually Matlab scripts (.m files).

It is possible to create tools for many other codes and software functionalities by creating xml descriptions of each specific tool. We can provide the xml schema for creating tools for computer savvy users. If communities would request wizards for a group of tools, for instance, tools created from GRASS package, then we might build additional wizards.

(1) Creating Command Line Tools

For this example, we will try to create a tool that provides directory listing given a directory location. The functionality can be tested by opening a cmd window by selecting Start->Run (on Windows machines), typing "cmd" and then typing "dir" in the cmd window. The result is the directory listing in the cmd window. The directory listing tool is a sequence of "cmd" and "dir" in the cmd window. The input is a string (text/plain) defining the directory location. The output is the text/plain that is generated in the cmd window.

To build the directory listing tool, start with the sequence File->New->External Tool Wizard. Figure 39 shows the first dialog that is designed for inserting a general description of the tool. Figure 40 through Figure 46 illustrate the meaning of the entries in each wizard dialog and how to populate them when the directory listing tool is created.

	External Tool Wizard	
Tool Inform	nation	Figur
Creates a ne	ew tool, all fields are required.	The
		first
Name:	cmd dirl	wiza
Version		
Description:	VI	dialog
Description:		aurii a the
		too
		creat
		on for
		direct
		ory
		listing
		Press
		"Next
		" to
		move
		to the
		secon
	< Back Next > Einish Cancel	wiza
		d

dialog.

훩 External Too	ol Wizard	<u> </u>
Command Line		
External tool comm	and line with options.	
Executable:	cmd	
🗆 Exitcode	0	
Capture stdout?	cmd dir output [text/plain]	
Options:	/C dir parameter(directory name [string] = "C:\tem	New Value
		New Parameter
		New Input
		New Output
		Remove
		Up
	▲	Down
cmd /C dir parame	eter(directory name [string] = "C:\temp") > cm	d dir output 🗦
	< Back Next > Finish	Cancel

Figure 40. The second wizard dialog for entering the executable and its parameters, inputs and outputs. The entry next to the label "Executable" refers to the name of the executable. If the "Exitcode" labeled check box is checked then the exit code will be verified to determine whether the application successfully completed its run or failed. The "Capture stdout" labeled check box is for enabling to pipe the output of execution in the execution window to a file. The buttons labeled as "New Value", "New Parameter", "New Input" and "New Output" are for inserting information about (a) the sequence of execution commands, (b) variables/parameters of the execution that should be directly modified before each run from Cyberintegrator, (c) inputs coming from data files already in Cyberintegrator or residing on disk, and (d) outputs generated by each run in the execution window or directly saved to a disk.

Command Line Option		×
Flag: /C		
Value: dir		
	ОК	Cancel

Figure 41. The dialog invoked from the second dialog shown in Figure 40 by clicking on the button "New Value". The entries correspond to the cmd command line options (cmd /C dir).

🋊 Comma	ndline Parameter 🛛 🗙
Flag:	
Name:	directory name
Description:	name of the directory to explore
Type:	string
Value:	C:\temp
🗆 Hidden?	
	OK Cancel

Figure 42. The dialog invoked from the second dialog shown in Figure 40 by clicking on the button "New Parameter". The entries correspond to the directory (value is set to C::\temp) that would be listed by default and is defined by a string that is passed to the command line.

훩 Data Cr	eation 🗙
Name:	cmd dir output
Description:	dir listing
Type:	text/plain
	text/plain
	image/ineg
	application/x-java-serialized-object ; class=

Figure 43. The dialog invoked from the second dialog shown in Figure 40 by checking the check box "Capture stdout?" and clicking on the button next to the check box. The entries correspond to the name, description and type of the output gathered by piping the text generated by "cmd /C dir" to a data set placed in the data pane of Cyberintegrator.

🛊 External Too	ol Wizard			<u>_ 🗆 ×</u>
External Resource	s			
List of external res	ources to bundle	with the tool.		
Files:				Add Delete
	< <u>B</u> ack	<u>N</u> ext >	Einish	Cancel

Figure 44. The third wizard dialog designed for adding external resources. For example, if an executable would depend on a set of dlls or jar files then one would click the button "Add" and select all additional files needed for the successful execution of the main executable. All external resources become parts of the context that is stored with the tool. If the tool would be shared then anyone might receive the tool with all external resources and run it assuming platform compatibility.

External To	ol Wizard			<u>_ 🗆 X</u>
A list of variables	set before the to	ol is executed.		
Name		Value Value		New
	< Back	Next >	Finish	Cancel

Figure 45. The fourth wizard dialog provided for creating environment variables needed to find external resources. For example, if an executable would depend on external resources that have to be specified by global environment variables then the variables could be set before the tool is executed.

Cyberintegrator			_ <i>_</i>
Cie Div Watche Geb			TR Charlenne
	(A a a a b a b a b a b a b a b b b b b b b b b b		E Gorneg
have filter that	Volume in drive C is IRM PRE 040	done Or bocomenic cino de objoc 35	<u> </u>
Tope firm text	Volume in drive C is IBM_RELOAD Volume Serial Marker is ICDA/GPO Directory of C1/temp 00/05/2007 12:46-164 008> 00/05/2007 12:46-164 008> 01/07/2007 16:46-164 008> 01/07/2007 06:04-164 128,243 12 01/07/2007 06:04-164 128,245 128 2 Dr(6) 32,214,737,408 bytes f	9940811881197 pgn 99498018901397 juan ggn Ree	
	Preview Annotations Tupelo View Attrib	butes	
D= Workflow D= Workflow Graph 🧟 Step Creator 🔅			- 0
General Information		* Parameters	
This section describes general information about the step that will be created.		This section describes parameters used by the step when being executed.	
Step name: cmd dr		directory name C:\temp	
Tool name: cmd dr		+ Inputs	
Execute Step		- Outputs	
		This section describes outputs generated by the step when being executed.	
		cua au colora l'auto ne colore.	14M of 2 <mark>1M</mark>
Cyberintegrator Ella Edit Mindau Mah			<u>_@</u> X
Can Die Weiter Cab			P2 70 chuddhar
(Down 12) Thomas (Down)	(D. Tank and do 12) (C. Law and and		Cyberintegr
and films for a	Annotations		<u></u>
		[August 1: 20172 5 (56, 55, 55, 55, 51)	
8 C plain			



Figure 46. Snapshots of Cyberintegrator after the directory listing tool was created (see the tool pane) and executed. The execution is accomplished by double clicking on the tool, specifying the name of the parameter (directory location) in the bottom pane and pressing the button "Execute Step" in the lower left part of the Step Creator window. After the step is completed, the data pane will show a new file called "cmd dir output [DONE]" while during the execution it would show "cmd dir output [RUN]". If the execution fails then it would be denoted by [FAIL]. The icons left to the name of the output also differ based on the status of

the execution. The top figure is obtained by right mouse clicking on the data set and selecting "Open in New Window". The right window (top figure) shows the directory listing for the C:\temp directory. The bottom figure is obtained by selecting the Tupelo View and it illustrates all metadata about the data set that was just created.

(2) Additional dialogs of interest:

There are several dialogs that have not been covered in the previous example and might be of potential interest to users. Two additional dialogs and their entries are described in Figure 47 and Figure 48.

ironment Variable Creation	×
JAVA_PATH	
C:\Program Files\Java\jre1.6.0_02\bin	
OK Cancel	
	IAVA_PATH C:\Program Files\Java\jre1.6.0_02\bin OK Cancel

Figure 47. The dialog invoked from the fourth wizard dialog shown in Figure 45 by clicking on the button "New". The entries correspond to the name of the environment variable and its value as illustrated for the case of setting the JAVA_PATH variable to point to the installation of java 1.6.

훩 Output D	ata 🗙
🔽 Pass on the	commandline?
Flag:	
Output:	N/A
E Filename?	
	OK Cancel

Figure 48. The dialog invoked from the second dialog shown in Figure 40 by clicking on the button "Output Data". The flag entry is designed for inserting any flags that would precede the name of the output file and would be parsed by the executable. The Filename checkbox and entry is for those executables that would have hard coded names of output files and we would them to be in the data pane after finishing the execution. The button next to the "Output" label

is invoking another dialog that allows describing output data placed in the data pane of Cyberintegrator immediately after launching the tool. The invoked dialog is identical with the dialog in Figure 43.

(3) Linking Tools:

In many cases, the tools have to be linked together the same way as scripts are created by forming a sequence of executions of multiple tools. In Cyberintegrator, linking can be accomplished by selecting a tool and passing the outputs of the previously executed tool to it. For instance, if the output of the directory listing tool would be passed to the Notepad tool then the output of the directory listing tool becomes the input to the Notepad tool. The linking is illustrated in Figure 49, Figure 50 and Figure 51. Figure 52 shows how to tag and sort resulting outputs.



Figure 49. Execution of the Notepad tool with input equal to the output from the directory listing tool. The Notepad tool is double clicked first in the Tool view. Then, the input file is selected to be the cmd dir output file in the Step Create View. The output of the notepad tool is denoted as cmd dir notepad output. Finally, the Notepad tool is executed by clicking on the "Execute Step" button.

🖡 cib15682.dat - Notepad	_ 🗆 🗙
File Edit Format View Help	
Volume in drive C is IBM_PRELOAD Volume serial Number is 1C2D-0F70	
Directory of C:\temp	
09/05/2007 12:46 PM <dir> . 09/05/2007 12:46 PM <dir></dir></dir>	
01/30/2007 06:04 PM 42,331 1169498616581397.pgm	
01/30/2007 06:04 PM 128,246 1169498616581397_plain.pgm 2 File(s) 170.577 bytes	
2 Dir(s) 3,214,737,408 bytes free	
	-

Figure 50. The directory listing followed by the Notepad tool execution leads to showing the directory listing in Notepad.



Figure 51. The workflow graph of the directory listing followed by the Notepad tool execution. (bottom pane)



Figure 52. Tagging data sets by CMD and Notepad tags allows sorting large collections of data sets. Tagging is enabled via the dialog (top figure) invoked by the combination of keys "ALT" and "T". Sorting is enabled via the "Sort by Tag" selection from the upper right corner of the Data view (the triangular shape icon) invoked by right mouse clicking. The result of sorting by tags is shown in the bottom figure.

Appendix ASoftware License

The software is owned by the University of Illinois at Urbana-Champaign and the appropriate licensing mechanism is currently under review.

Appendix BKnown Issues

There is a list of known issues with the current alpha release that are recorded in the NCSA JIRA system at

http://torcida.ncsa.uiuc.edu:8080/jira/secure/IssueNavigator.jspa?reset=true&mode=hide&sort er/order=DESC&sorter/field=priority&resolution=-1&pid=10052&fixfor=10046

JIRA is a bug tracking application where the development team and the users and testers can report software bugs. One can monitor the progress of the Cyberintegrator project by going to the NCSA JIRA web site at <u>http://torcida.ncsa.uiuc.edu:8080/jira/browse/CBI</u>. This site should be also used for filing a new software bug as illustrated in Figure 53.

Browse Project - NCSA Grid Computing En	vironements JIRA - Micro	soft Internet Explorer			X
Eile Edit View Favorites Icols Help					2
😋 Back 🔹 🐑 💌 😰 🐔 🔎 Search 👷 Favo	rites 🕢 😥 🗟 • 😓 🖩 • 🗖	J 01• 🋍 😐 🕲 🕉			
Address 🕘 http://torcida.ncsa.uiuc.edu:8080/jira/browse/0	BI				💌 🔊 வ 💽 💌
⊻! • <i>e</i> •	web 🛛 🖉 🕞 • 🕘 • 🚇 🗛 U	pgrade your Toolbar Now! • 🖂 Mail • 🕲 My Yahoo! •	🐑 Baseball 🔹 🥥 Games 🔹 🦓 Music	• 💭 Answers 🔹 💜 Personals	• De Sign In • Vinke
🕒 Browse Project - NCSA Gri 💄 Browse Project - NCSA	Sri 📄 ISDA NCSA UIUC.EDU DO	Add Tab			
Google G vsystem download ▼ Go + ⊘ 😒 🗗 v 😭	Bookmarks 🛛 💝 Check 💌 👫 Trar	slate 🔻 🔨 AutoLink 🔻 🍺 Send to 🛛 🌽 🖳 JIRA 🖳 bug	💽 tracking 🗟 system 🗟 download		Settings Ienovo
HONE BROWSE PROJECT FIND ISSUES SREATE NEWS	ISSNE				DUTCK SEARCH:
All Projects : Cyberintegrator (Key: CBI)					
Project Las Expl.Boose D'anstan Join Distanza Join Status distanza Status distanza Status distanza Status distanza Read Marc Change Lee Property for Component/s Read Marc Change Lee Property for Vel Commission in such component) Read Status Read Status	001 10 2 3 69 Powered by <u>references 2004</u>	Versions (with open issues due to be fixed per version) 2.2 2.2 2.2 2.2 2.2 2.2 2.2 2.	64 20	Reports Trans Tracking Resord Variation State (Resord Variantia and Resord Variantia and Resord Preset Filter • Unattracked • Unattracked • Institution • Institution	- Emothed recentr - 0.400d r
					-
Done					internet

Figure 53. The NCSA JIRA web page for Cyberintegrator at

<u>http://torcida.ncsa.uiuc.edu:8080/jira/browse</u>. One can report a bug by clicking on the "create new issue" tab. and proceeding through a set of windows to describe the software bug.