MIND STEP

MODELLING INDIVIDUAL DECISIONS TO SUPPORT THE EUROPEAN POLICIES RELATED TO AGRICULTURE

Deliverable D 7.5: Setup of the version control system and the continuous integration tool

AUTHORS	Albert Brouwer (IIASA)
APPROVED BY WP MANAGER:	Gocht, Alexander (Thuenen)
DATE OF APPROVAL:	06-07-2021
APPROVED BY PROJECT COORDINATOR:	Hans van Meijl (WR)
DATE OF APPROVAL:	09-07-2021
WP NUMBER & TITLE	WP 7 ICT platform of MIND STEP and the MIND STEP model toolbox
DISSIMINATION LEVEL	Public
PROJECT WEB SITE:	https://mind-step.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

TABLE OF CONTENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

EXECUTIVE SUMMARY

To determine version control system (VCS) and continuous integration (CI) requirements, MIND STEP partners were polled by means of a questionnaire. The outcome showed a predominant use of Subversion as a VCS and Windows as a processing platform. The results and options were summarized in a concept note, and after discussion it was decided that IIASA provide a Subversion server for VCS and a Jenkins server for CI. Both servers are now available for MIND STEP use.

GitLab was discussed as an additional option that would provide Git repository hosting for VCS, and an alternative CI workflow with rich development collaboration features. Motivated by the MIND STEP requirements, the IIASA ICT department also set up a GitLab server with account registration open to MIND STEP partners.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

1. BACKGROUND

Collaborative development of models greatly benefits from version control of code and data. A VCS keeps track of who did what and when to the various files and directories comprising a model. To allow partner collaboration, the VSC must be internet accessible.

Automated testing, integration, and deployment requires infrastructure that can run models, or small parts thereof. CI systems provide both a platform for processing as well means to organize the processing flow as so-called pipelines. The CI system is enabled to access the VSC system(s) and retrieve the files required for execution of a particular pipeline. Partner collaboration on the pipelines required the CI system to be internet accessible.

To guarantee privacy of data, cloud-based solution for version control storage and CI processing was considered to not be an option. Hence, on-IIASA-premises hosting of the VCS and CI systems was a precondition.

2. QUESTIONNAIRE

The questionnaire polled partners for technical details of their models and development practices to thereby constrain VCS and CI selection, deployment, and configuration. Queries included the size of the overall code and data as well as the largest file, operating system and language runtimes in use, dependencies such as packages and solvers, and details of existing test practices.

Notable outcomes included the almost exclusive use of Windows, the use of GAMS by two thirds of the models, and R by about half. Subversion was used by most teams for version control. Git, though being more modern and capable, was used by only two teams. In terms of dependencies, numerous open-source packages were in use and CONOPT and CPLEX were the dominant GAMS solvers.

Test scripts were reported to be implemented for about half the models. Making available the MIND STEP infrastructure is expected to result in expansion of the test automation and coverage for the participating, thereby improving quality. The capabilities of the CI systems (Jenkins and GitLab) encourage a more granular and platform-agnostic approach to testing.

A concept note containing the outcome of the questionnaire and summarizing the options for VCS and CI systems was elaborated for partner review. After presentation and discussion, Subversion was chosen as the VCS system, and Jenkins as the CI system. The deployment of a MIND STEP Subversion and Jenkins server at IIASA was initiated. In addition, the IIASA ICT department was informed that a user-friendly way of hosting Git repositories that are Internet-accessible would be helpful for MIND STEP. After an elaborate discussion of options, GitLab was chosen for filling the latter need.

The questionnaire questions and responses are tabulated in Annex II.

3. MIND STEP VCS AND CI SERVERS

At the end of 2020, IIASA ICT delivered the Subversion VCS and an Jenkins CI servers dedicated to MIND STEP. In addition, early 2021, a GitLab server for IIASA internal and partner use was deployed. The email domains of MIND STEP partners were whitelisted for GitLab account creation. The GitLab server provides for both Git VCS repository hosting as well as CI pipeline execution.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

The Subversion server is administered by IIASA ICT, with support via an email address dedicated to MIND STEP. The Jenkins and GitLab servers have a self-service web interface that allows partners to create and configure CI jobs and, in the case of GitLab, Git repositories. The GitLab Git repositories can be made public or private and associated with a wiki and issue tracker for collaborative documentation and development. A

To guide partners with acquiring access and basic use of the toolbox, a website was created to share details on access procedures with MIND STEP partners. Links from the table of contents are reproduced below:

- Getting started
- Access
    - Subversion server
    - Jenkins server
    - GitLab server
- Privacy
- Basic use
    - Subversion
    - Jenkins
    - GitLab

Note in particular the GitLab section mentioning the availability of a MIND STEP group.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

ACKNOWLEDGMENT & DISCLAIMER

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817566. Neither the European Commission nor any person acting on behalf of the Commission is responsible for how the following information is used. The views expressed in this publication are the sole responsibility of the author and do not necessarily reflect the views of the European Commission.

ANNEX I MIND STEP WP7 TEAM

Mr Sebastian Neuenfeldt (THUENEN)
Mr Alexander Gocht (THUENEN)
Mr John Helming & Mr Marc Müller (WEcR)
Mr Albert Brouwer, Mr Stefan Frank, Mr Ian McCallum, Mr Petr Havlik (IIASA)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

ANNEX II QUESTIONNAIRE QUESTIONS AND RESPONSES

Model name	Technical contact	Code size	Data size	Largest data file	OS	Language runtime(s)	Python packages
Generic modular farm level model	Marc Müller	15 MB		85 MB	Windows	GAMS 28.2+, R 3.6.1, Python 3.7+, JRE 8+	os, gams, numpy, pandas, seaborn, matplotlib, scipy, sklearn.decomposition import PCA, shapefile, geopandas as gpd, descartes, shapely, sqlite
GHG model	Marc Müller	15 MB	750 MB	85 MB	Windows	GAMS 28.2+, R 3.6.1, Python 3.7+, JRE 8+	gams, numpy, pandas, seaborn, matplotlib, scipy, sklearn.decomposition import PCA, shapefile, geopandas as gpd, descartes, shapely, sqlite
IFM-CAP	Dimitrios Kremmydas	4 MB	4 GB (including data generated for scenario run)	1.3 GB	Windows	GAMS 28.2+, R 3.6+	
FarmAgripolis	Franziska Appel			150 kB	Windows	compiled C++, Qt for GUI	
Supply Chain Model	Paolo Sckokai, Martin Banse	small		< 10 MB	Windows	Stata, R	
Risk behaviour module	Frank Offerman, Paolo Sckokai	small		<1 MB unless weather data is included	Windows	probably GAMS	
FarmDyn	David Schaefer	240 MB		64 kB	Windows	GAMS 29.0+, R 3.2.1, GAMS-embedded Python, JRE for GUI	
MAGNET	John Helming	21 GB combined	21 GB combined	300 MB	Windows	GEMPACK (license and GTAP data base license required), GEMPACK, DSS, GTree	
AgriSpace	Klaus Mittenzwei	4.2 GB combined	4.2 GB combined	150 MB	Windows	GAMS	
Farm Exit Model	Sebastian Neuenfeldt, Alexander Gocht	1 MB	1 GB	800 MB	Windows	R 3.5.0+	
Crop model	Fabienne Femenia				Windows	R 3.51+	
GLOBIOM	Albert Brouwer	8.4 GB combined	8.4 GB combined	932 MB	Windows/ Linux	GAMS 29.1+, R3.5+	

Model name	R packages	GAMS solvers	VCS in use for model	VCS/store in use for data	Test scripts implemented?	Language runtime(s) for tests	Packages required for tests
Generic modular farm level model	lhs, gdxrrw , mc2d, mvtnorm, Matrix, gclus, reshape2, ggplot2, plyr	CPLEX, CONOPT	SVN	SVN, gdx, csv, xlsx, sqlite	yes	GAMS 28.2+, R 3.6.1, Python 3.7+, JRE 8+	
GHG model	lhs, gdxrrw, mc2d, mvtnorm, Matrix, gclus, reshape2, ggplot2, plyr	CPLEX, CONOPT	SVN	SVN, gdx, csv, xlsx, sqlite	yes	GAMS 28.2+, R 3.6.1, Python 3.7+, JRE 8+	
IFM-CAP	fadnUtils (in-house package to handle FADN data), data.table, reshape2, ggplot2	CONOPT, SBB	SVN	SVN	yes	GAMS	
FarmAgripolis		GLPK solver, GPL license	Git	Git	no		
Supply Chain Model		CONOPT, SBB					
Risk behaviour module		depends on the model to which the risk module will be linked		Postgres SQL, gdx for output			
FarmDyn	lhs, gdxrrw, mc2d, mvtnorm, Matrix, gclus, reshape2	CPLEX (recommended) or GUROBI	SVN	SVN	yes	same as model	same as model
MAGNET			SVN	SVN	no		
AgriSpace		PATH or any other MCP-solver	SVN	SVN	no	GAMS	
Farm Exit Model	probably: data.table, tidyverse, felm, plm, hurdle		SVN	N/A: confidential data for Germany, processing has to be at a dedicated facility	no		
Crop model							
GLOBIOM	tidyverse, gdxrrw	CPLEX, CONOPT	SVN, Git	SVN, Git, csv, xlsx, gdx	yes	GAMS, R	tidyverse, gdxxrw, testthat

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

CONSORTIUM DESCRIPTION

The consortium of MIND STEP consists of 11 partners from 7 countries in Europe (the Netherlands, Germany, Austria (IIASA), Italy, France, Spain (JRC-Seville), Norway and Hungary). It includes partners from the private and public sector representing:

- Academia and higher education (UBO, UCSC, WU).
- SME dealing with research consultancy, data collection, strategic advice, normalization and policy in the field of energy, environment and sustainable development. This SME has also a strong track record in the field of communication, stakeholder engagement and exploitation (GEO)
- Public government bodies dealing with agricultural and environmental research and data collection and building agricultural models at different scales (WR, IIASA, IAMO, THUENEN, INRA, NIBIO, JRC)

The consortium has been carefully constructed in such a way that it is capable of jointly managing all activities and risks involved in all project stages. Each partner contributes its own particular skills, (inter) nationally wide network and expertise, and has a critical role in MIND STEP. Partner expertise smoothly complements each other and all together form the full set of capabilities necessary to lead MIND STEP to a success. Achieving the overall objective is determined by all partners in the consortium as well as their ability to involve other interested stakeholders in the process of developing, validating and disseminating the IDM models, indicators and methodologies (WR, UBO, IAMO, UCSC, WU, THUENEN and INRA) and linking IDM models to current agricultural policy models (WR, IIASA, UBO) included in the MIND STEP model toolbox. Dissemination and communication activities are steered by partner GEO who has graphic design, IT and marketing communication teams to deliver out-of-the-box and novel solutions for dissemination and communication and JRC who has a large network with policy makers. GEO has experience in leading comparable activities in H2020 projects as UNISECO and COASTAL. The coordinator WEcR is part of Stichting Wageningen Research (Wageningen Research Foundation, WR). WR consists of a number specialised institutes for applied research in the domain of healthy food and living environment. WR collaborates with Wageningen University (WU) under the external brand name Wageningen University & Research. One of the strengths of Wageningen University & Research (including WR) is that its structure facilitates and encourages close cooperation between different disciplines. The institutes Wageningen Economic Research (proposed coordinator of MIND STEP, WEcR) and Wageningen Environmental Research (WEnR) are involved in this proposal. The One-Wageningen approach will also be applied to MIND STEP. WEcR has a long standing reputation of leading large scale EU projects, such as SUPREMA, Foodsecure, SUSFANS, FLINT, SAT-BBE, and SIM4NEXUS.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.