

ИНТЕЛИГЕНТЕН КОМПОНЕНТ ЗА ПРЕДУПРЕЖДЕНИЕ ЗА КОНТРОЛНИ ПРОВЕРКИ В СИСТЕМАТА GENEVANK

Пенчо Малинов, Емил Дойчев,
Себиха Маданска, Ася Стоянова-Дойчева

***Резюме.** Разработването на система за управление на растителните генетични ресурси е с цел повишаване на сигурността за запазване на генофонда в националната генбанка в България, която се намира в Института по растителни генетични ресурси „Константин Малков“ – Садово. Към момента в нея се съхраняват над 60 000 образци от различни видове. Основен компонент на системата осъществява управление на контролни проверки за кълняемост на семената с цел запазването им за възпроизводство. Статията представя модул свързан с управление на контролните проверки, които се провеждат от лаборантите в генбанката. Разгледана е архитектурата на системата, разработена до момента и включването на интелигентен агент за проследяване и предупреждение при нужда от контролни проверки.*

Ключови думи: интелигентни агенти, растителни генетични ресурси, генбанка.

Въведение

Съхранението и опазването на растителния генофонд на България е от основно национално значение. В Института по растителни генетични ресурси (ИРГР) в град Садово се намира най-голямата генбанка в България. В нея се съхраняват над 60 000 образци на растителни генетични ресурси, голяма част, от които са на местни видове и сортове. От 2020 година екип учени от ИРГР съвместно с колеги от Пловдивския университет и Института по информационни и комуникационни технологии на БАН работят по проект на фонд „Научни изследвания“ за създаване на информационна система за управление на растителния генофонд в генбанката. Системата има за цел да улесни работата на служителите в института и да подобри качеството на работата им.

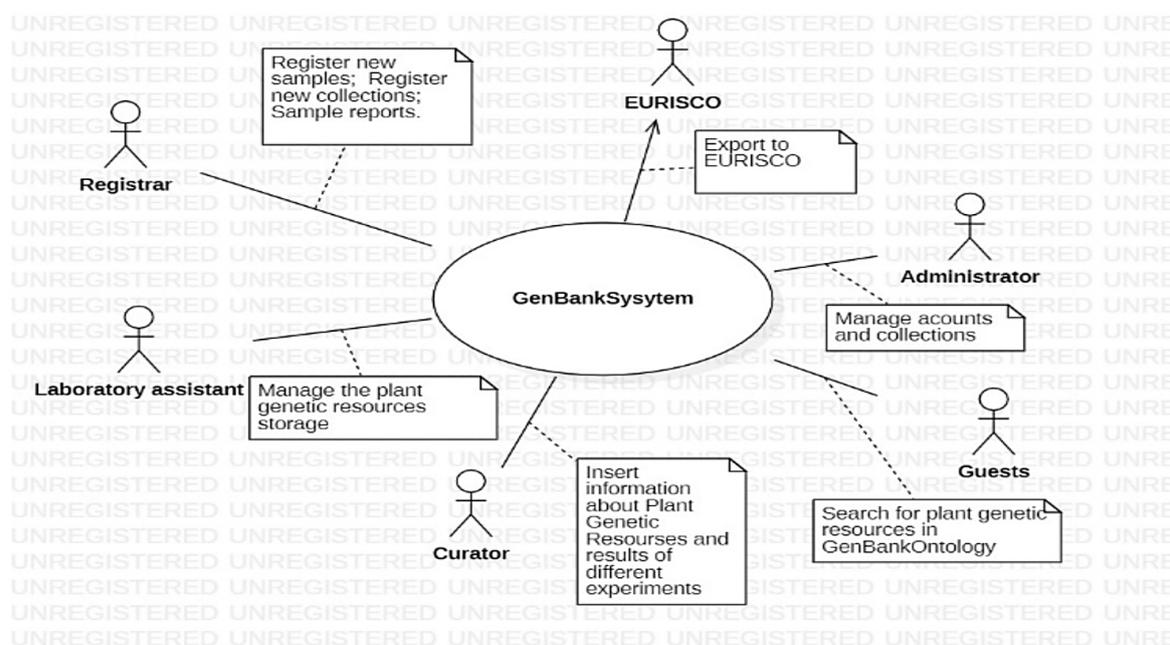
Основните изисквания към системата са свързани със стандарта EURISCO [1], използван за унифициране на информацията, която описва растителния генофонд. Всеки образец се описва с паспортни данни, които са в иден-

тичен формат за целия Европейски съюз. От друга страна, международният договор за растителни генетични ресурси [2] и протоколът от Нагоя, Япония [3] изискват данните за различните генетични ресурси да бъдат достъпни за всички заинтересовани страни – институти, университети, фермери и други. Това задължава системата да предостави модул, чрез който тези заинтересовани страни да имат достъп до информацията за растителните генетични ресурси. За тази цел беше разработена *GenBankOntology* [4]. Тя съдържа допълнителна информация за образците в генбанката, която позволява достъп до информацията включително на потребители, които не са специалисти в областта.

Основната цел на статията е да покаже разширение на архитектурата на системата за управление на растителните генетични ресурси *GeneBankSystem* [5] с интелигентни компоненти, които улеснят и подобрят работата в генбанката, фокусирайки се върху интелигентен компонент за предупреждение за контролни проверки на образците, съхранявани в нея.

Функционалност и архитектура на системата за управление на растителните генетични ресурси

Функционалностите на *GeneBankSystem* са свързани с нуждите на нейните потребители. Това са различни специалисти и учени, служители на ИРГР-Садово. Фигура 1 представя use-case диаграма на основните потребители на системата.



Фигура 1. Функционалности на *GeneBankSystem*

Основно задължение на служителите в генбанката на ИРГР-Садово

е осъществяването на регистрацията и управление на всички растителни генетични ресурси (РГР) постъпващи в института. Данни за тези ресурси се въвеждат в системата като индивидуални образци, всеки от които представя РГР с определен род, вид и година на размножаване. Паспортните данни на образците, дефинирани в съответствие със стандарта EURISCO, осигуряват единна и стандартизирана информация за техните характеристики. Всеки образец се разпределя в една или няколко колекции на генбанката. Колекциите за съхранение предоставят различни режими на съхранение – краткосрочно, средносрочно и дългосрочно съхранение, които гарантират подходящи условия за опазване на генетичния материал. Допълнително се поддържат и колекции на куратори, които съдържат единствено образци за краткосрочно съхранение и се използват за провеждане на различни експериментални дейности.

Работният процес в генбанката се реализира от няколко групи потребители на системата: *Регистратор*, *Администратор*, *Лаборант*, *Куратор*, *външната система EURISCO* и *Гост*. Всеки от тези потребители изпълнява специфични функции в рамките на управлението на РГР.

Кураторите провеждат експерименти с образците от своите колекции, като документират и анализират резултатите в системата. Регистраторът отговаря за въвеждането на новопостъпващи образци, включително за записването на техните паспортни данни и за разпределянето им в подходящи колекции за съхранение и в колекциите на куратори. Лаборантите поддържат колекциите за съхранение и управляват информацията, свързана със срока на годност на образците, качеството на семената, проведените експерименти и получените резултати.

Важен нормативен ангажимент на всяка генбанка е регистрирането на наличните образци в системата EURISCO. В този контекст *GeneBankSystem* предоставя функционалност за експорт на паспортните данни на образците във формат, съвместим с изискванията на EURISCO, което осигурява интеграция и обмен на информация на международно ниво.

Достъпът на външни потребители до системата се осъществява чрез ролята *Гост*. Тези потребители могат да преглеждат информация за растителните генетични ресурси единствено чрез разработената в системата онтология, което гарантира контролирано и структурирано предоставяне на данни.

Архитектурата на системата се състои от три основни слоя:

- *Front-End Layer* – използва се за визуализиране на информация за растителните генетични ресурси. Интерфейсът е организиран спрямо

достъпа на потребителя. Информацията за растителните генетични ресурси е свободно достъпна от нерегистрирани потребители и за тази цел компонентите на интерфейса използват *GenBankOntology*.

- *Back-End Layer* – състои се от три основни компонента – *GenBankServices*, *GenBankAssistant* и *GenBankSynchAssistant*. *GenBankServices* е компонент, който реализира основните функционалности за съхранение и обработка на образци от растителните генетични ресурси. Информацията, която се въвежда и използва от служителите, се записва в база данни – *GenBankDatabase*. Компонентите *GenBankAssistant* и *GenBankSynchAssistant* са разработени като оперативни асистенти. *GenBankAssistant* е създаден за обработка и предоставяне на информация за растителните генетични ресурси, използвайки *GenBankOntology*. Основната му функционалност е да търси растителни генетични ресурси в онтологията според специфични критерии. *GenBankSynchAssistant* е разработен за синхронизиране на базата данни и онтологията в приложението. За тази цел този оперативен асистент следи състоянието на базата данни и при въвеждане на нов образец се уверява, че информацията за него е записана в онтологията.
- *Knowledge Base Layer* – този слой съдържа знания за растителните генетични ресурси. Разработени са два основни компонента – *GenBankDatabase* и *GenBankOntology*. Базата данни съдържа цялата информация за растителните генетични ресурси, както и всички експерименти и резултати, проведени върху тях. Достъпът до базата данни е ограничен за външни потребители. Онтологията за растителни генетични ресурси – *GenBankOntology*, е предоставена за свободен достъп. Тя съдържа информация, достъпна за всички, и също така позволява интелигентно търсене от заинтересованите страни.

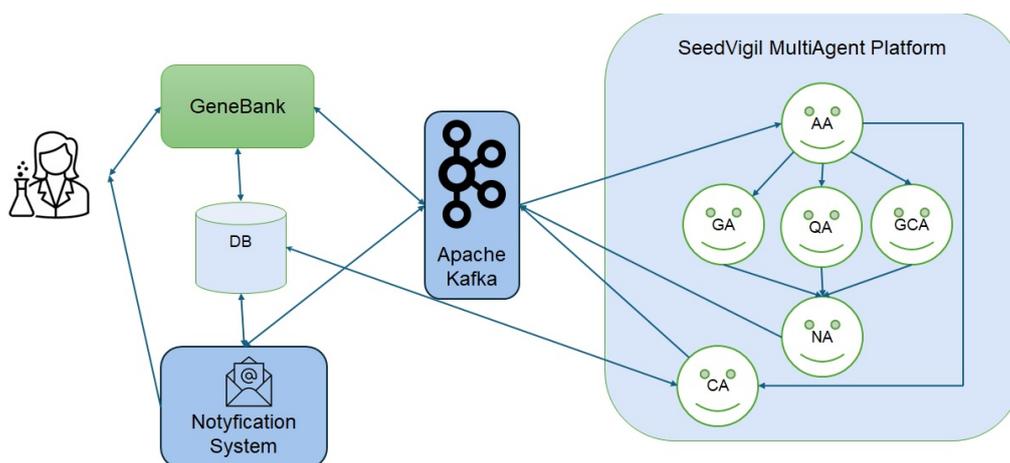
Интелигентен компонент за предупреждение за контролни проверки в GeneBank системата

Компонентът за управление на контролните проверки в *GeneBank* системата се използва от лаборантите на генбанката. Техните основни задачи са свързани с регистрацията на репродукции в генната банка, определяне на вида съхранение, провеждане на тестове за кълняемост на образци за дългосрочна и средносрочна колекция и прехвърляне на семена от дългосрочна или средносрочна в работна колекция с цел репродукция.

Основните проблеми, които възникват в работата на лаборантите са свързани с непрекъснатото проследяване на регистрациите на образците в генбанката с цел определяне на тези от тях, които трябва да бъдат подложени

ни на контролни проверки. За всеки тип съхранение има различни критерии, които определят дали образецът или репродукцията трябва да бъде подложена на контролна проверка. В *GeneBank* системата информацията за тези образци се извежда на екрана на лаборанта след филтриране по определени критерии. Това може да доведе до човешка грешка и пропускане на контролни проверки на някои от образците, а в най-лошия случай и до загуба на образеца. За решаване на този проблем решихме да разработим интелигентен компонент, който да предупреждава лаборантите за наличие на образци, които трябва да бъдат подложени на контролни проверки, а също така да алармира при настъпили аномалии в характеристиките, свързани с измерванията от контролните проверки – кълняемост, кълняема енергия, влажна маса, брой непокълннали семена, твърди и мъртви семена. Когато някой от праговете на тези характеристики бъде нарушен, тази информация незабавно ще бъде изпратена до лаборантите в генбанката.

Предложена е мулти-агентна архитектура (Фигура 2), базирана на *Belief-Desire-Intention (BDI)* архитектура [6]. Тя предоставя мощен модел за изграждане на интелигентни агенти, способни на самостоятелно вземане на решения посредством вътрешни убеждения, желания и планирани намерения. В контекста на системата за предупреждение за контролни проверки в генбанката в ИРГР-Садово, интеграцията на BDI агентите позволява гъвкав и адаптивен анализ на големи обеми данни, получени от системата *GeneBank* и предавани чрез *Apache Kafka* [7].



Фигура 2. Архитектура на интелигентния компонент за предупреждения за контролни проверки

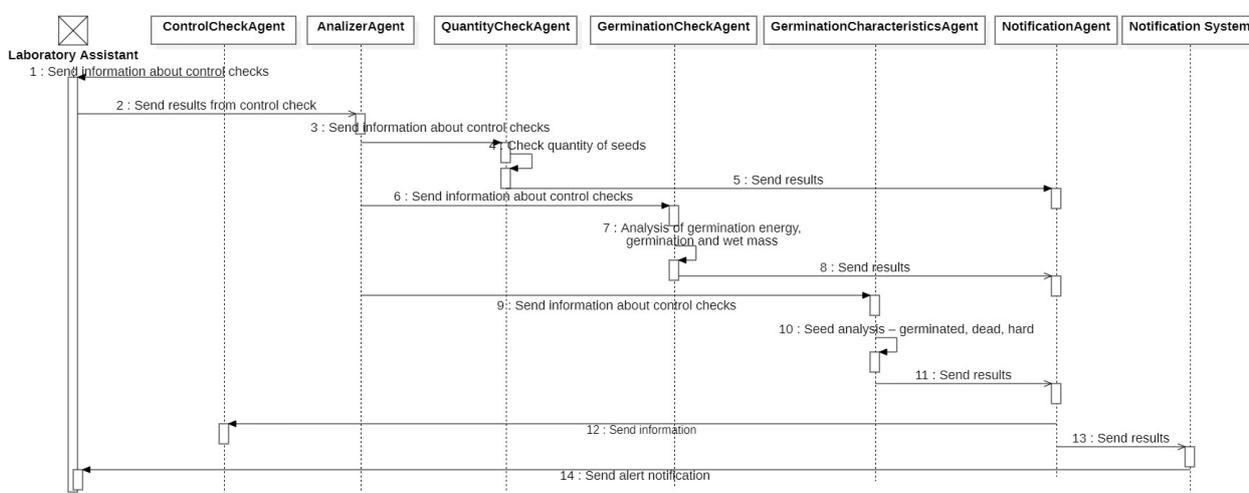
Архитектурата се състои от следните компоненти:

- Система *GeneBank* – източник на данни за контролни проверки;
- *Kafka Broker* – среда за публикуване и абониране за съобщения;

- *ControlCheckAgent (CA)* – проверка за необходимост от контролни проверки. Има директен достъп до базата данни на генбанката;
- *Analyzer Agent (AA)* – първоначален анализатор. Получава данни от Kafka;
- *Quantity Check Agent (QA)* – анализира наличните количества на семена;
- *Germination Check Agent (GA)* – анализира енергията на покълване, покълването и влажната маса;
- *Germination Characteristics Agent (GCA)* – анализира характеристиките на семената – покълнали, мъртви, твърди;
- *Notification Agent (NA)* – известява и публикува аларми в Kafka;
- *Notification System* – получава аларми и изпраща имейли до лаборантите.

Комуникацията между агентите се осъществява на базата на стандарта FIPA ACL [8]. Той осигурява структурирана и формална комуникация, която улеснява автоматизираното разбиране и изпълнение на съобщенията. Поддържа сложни взаимодействия и преговори между агентите, важни за кординиране на комплексни действия. В комбинация с BDI модела, FIPA ACL позволява на агентите да изпълняват динамични планове и да адаптират поведението си според текущата ситуация като улеснява интероперабилността с други системи и интеграция на агенти, реализирани на различни платформи.

Сценарият за комуникация между агентите е представен на Фигура 3.



Фигура 3. Комуникация между агентите

ControlCheckAgent изпраща съобщение до лаборанта за наличие на образци за контролни проверки. След въвеждане на данните от контролна проверка в *GeneBank* системата, *Analyzer Agent* получава данните от Kafka и след анализ изпраща данните до *Quantity Check Agent*, *Germination Check Agent* и *Germination Characteristics Agent*. Всеки от тях изпълнява проверка на характеристиките, за които е отговорен – количество семена след контролна проверка, характеристики на кълняемостта и характеристики на семената. Агентите изпращат резултата получен от съответната проверка до *Notification Agent*, който от своя страна изпраща съобщение до *ControlCheckAgent* и *Notification System*. *ControlCheckAgent* има задачата да промени информацията за контролни проверки в дашборда на лаборанта, а *Notification System* да информира за резултатите от контролната проверка чрез изпращане на имейли до лаборантите.

За реализацията на агентите се използва Jadex [9]. Това е модулна, отворена платформа за разработка на интелигентни BDI агенти. Тя поддържа структуриран дизайн на агенти, включващ дефиниране на убеждения (*Beliefs*), желания (*Desires*) и намерения (*Intentions*), както и планиране и изпълнение на действия, което я прави изключително подходяща за системата за предупреждения при контролни проверки в генбанката. Този модел отразява най-добрите практики за интеграция на Kafka с мулти-агентни системи и акцентира върху мащабируемостта, координацията и надеждността на системата, съобразно съвременните изисквания за такива решения.

Заклучение

В заключение, разработената система значително улеснява работата на лаборантите в генбанката, като оптимизира процесите по наблюдение и управление на състоянието на образците. Чрез автоматизираното проследяване се намалява вероятността от пропускане на навременни контроли на семената, което в противен случай може да доведе до влошаване на качеството на образците или дори до тяхната загуба. Следва да се подчертае, че към момента в генбанката се съхраняват 66 399 образци, като броят им продължава да нараства, което допълнително обосновава необходимостта от надеждни и ефективни инструменти за тяхното управление.

Благодарности

Това изследване е финансирано по проект КП-06-Н86/9/09.12.2024 „IS-PGR-SADOVO Интелигентна система за управление на българския растителен генофонд, съхранен в генбанката на ИРГР-Садово Фундаментални изследователски проекти на Националния научен фонд на Министерството на

образованието и науката на Република България.

Литература

- [1] EURISCO, <http://eurisco.ecpgr.org>
- [2] ITPGRFA, International Treaty on Plant Genetic Resources for Food and Agriculture, Rome, Italy, 2009
- [3] CBD, 2011, Nagoya protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilisation to the Convention on Biological Diversity. United Nations Environmental Programme.
- [4] A. Stoyanova-Doycheva, V. Ivanova, E. Doychev, K. Spassova, Development of an Ontology in Plant Genetic Resources, *2020 IEEE 10th International Conference on Intelligent Systems (IS)*, Varna, Bulgaria, 2020, pp. 246–251, doi:10.1109/IS48319.2020.9199935
- [5] E. Doychev, P. Malinov, N. Velcheva, Z. Ducheв, A Genebank Architecture: A Distributed System for Management of Plant Genetic Resources, *2020 IEEE 10th International Conference on Intelligent Systems (IS)*, Varna, Bulgaria, 2020, pp. 580–583, doi:10.1109/IS48319.2020.9199972
- [6] R. Bordini, M. Wooldridge, J. Hubner, *Programming Multi-Agent Systems in AgentSpeak using Jason*, John Wiley & Sons, 2007
- [7] Apache Kafka, <https://kafka.apache.org/>
- [8] FIPA ACL standard, <https://site.ieee.org/pes-mas/agent-technology/standards-and-interoperability/#:~:text=The%20FIPA%20Agent%20Communication%20Language,the%20communicative%20act%20it%20performs>
- [9] Jadex, <https://download.actoron.com/docs/releases/jadex-3.0.98/jadex-mkdocs/tutorials/bdiv3/02%20Your%20first%20BDI%20Agent/>

Пенчо Малинов¹, Емил Дойчев¹,

Себиха Маданска¹, Ася Стоянова-Дойчева¹

¹ Пловдивски университет „Паисий Хилендарски“

Факултет по Математика и Информатика

Бул. „България“ № 236, 4027, Пловдив, България

Автор за кореспонденция: astoyanova@uni-plovdiv.bg

INTELLIGENT WARNING COMPONENT FOR CONTROL CHECKS IN THE GENE BANK SYSTEM

Pencho Malinov, Emil Doychev,
Sebiha Madanska, Asya Stoyanova-Doycheva

Abstract. *The development of a system for managing plant genetic resources is aimed at increasing the security of preserving the gene pool in the national gene bank in Bulgaria, which is located in the Institute of Plant Genetic Resources in the town of Sadovo. Currently, over 60,000 samples of various species are stored in it. A main component of the system is related to the management of control checks for seed germination in order to preserve them for reproduction. The article presents the module related to the management of control checks, which are carried out by laboratory technicians in the gene bank. The architecture of the system developed to date and the inclusion of an intelligent agent for tracking and warning when control checks are needed are examined.*

Key words: Intelligent Agents, Plant Genetic Resources, Gene Bank.

Acknowledgments

This research was funded under project KP-06-N86/9/09.12.2024 "IS-PGR-SADOVO Intelligent system for management of the Bulgarian plant gene pool stored in the gene bank of IRGR-Sadovo", Fundamental Research Projects of the National Science Fund of the Ministry of Education and Science of the Republic of Bulgaria.