

Leveraging AI and Human Expertise in Exotic Pest Early Warning & Pest Information Systems



Godshen Robert

NC STATE UNIVERSITY

Center for Integrated Pest Management

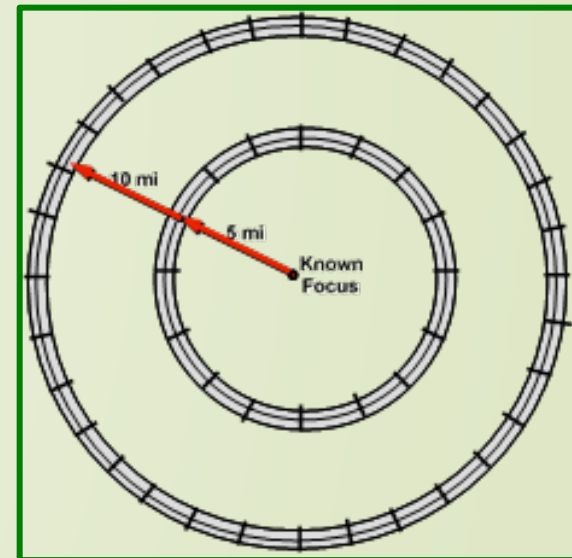
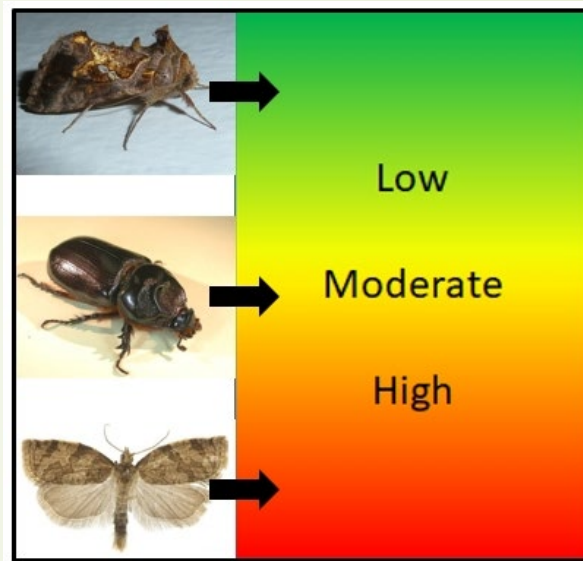
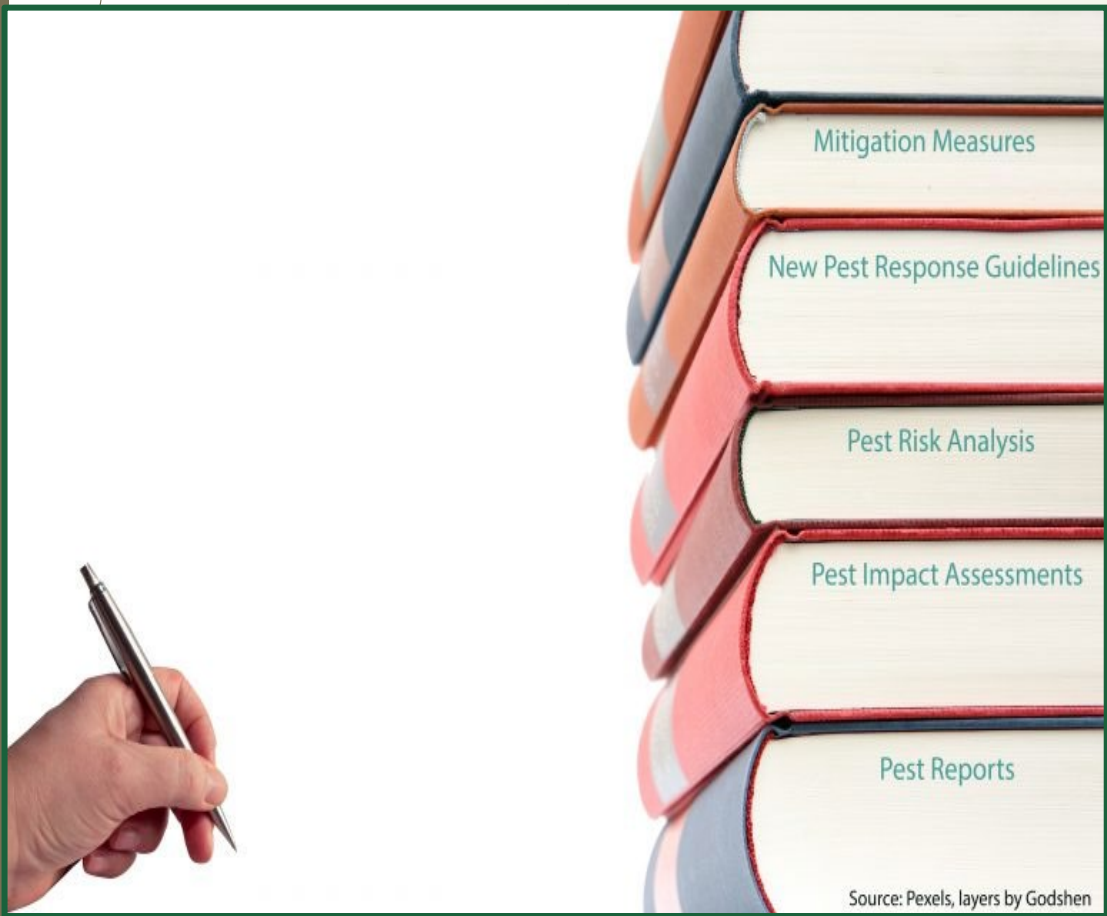
Center for Integrated Pest Management (CIPM)





70-80
employees

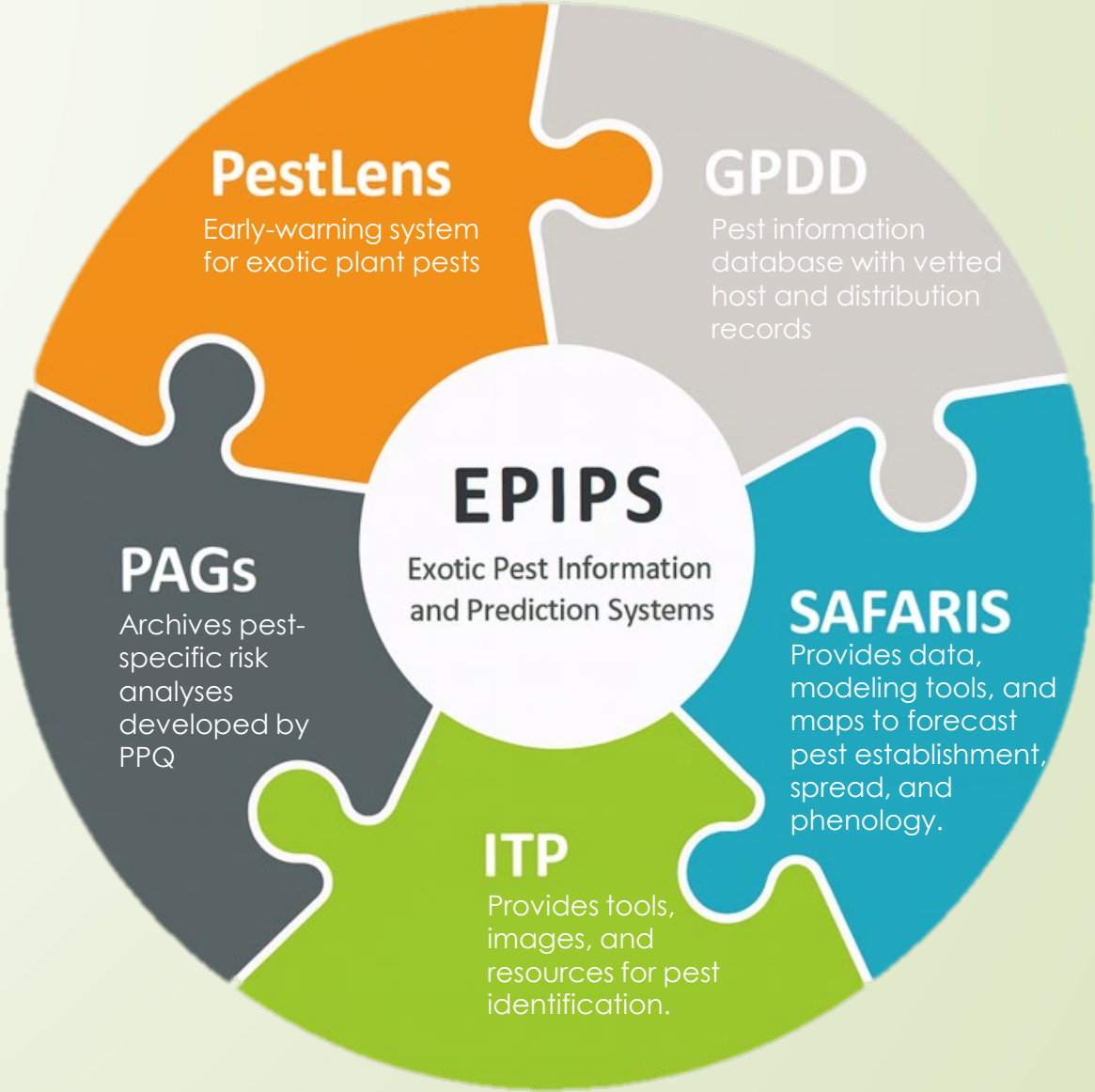


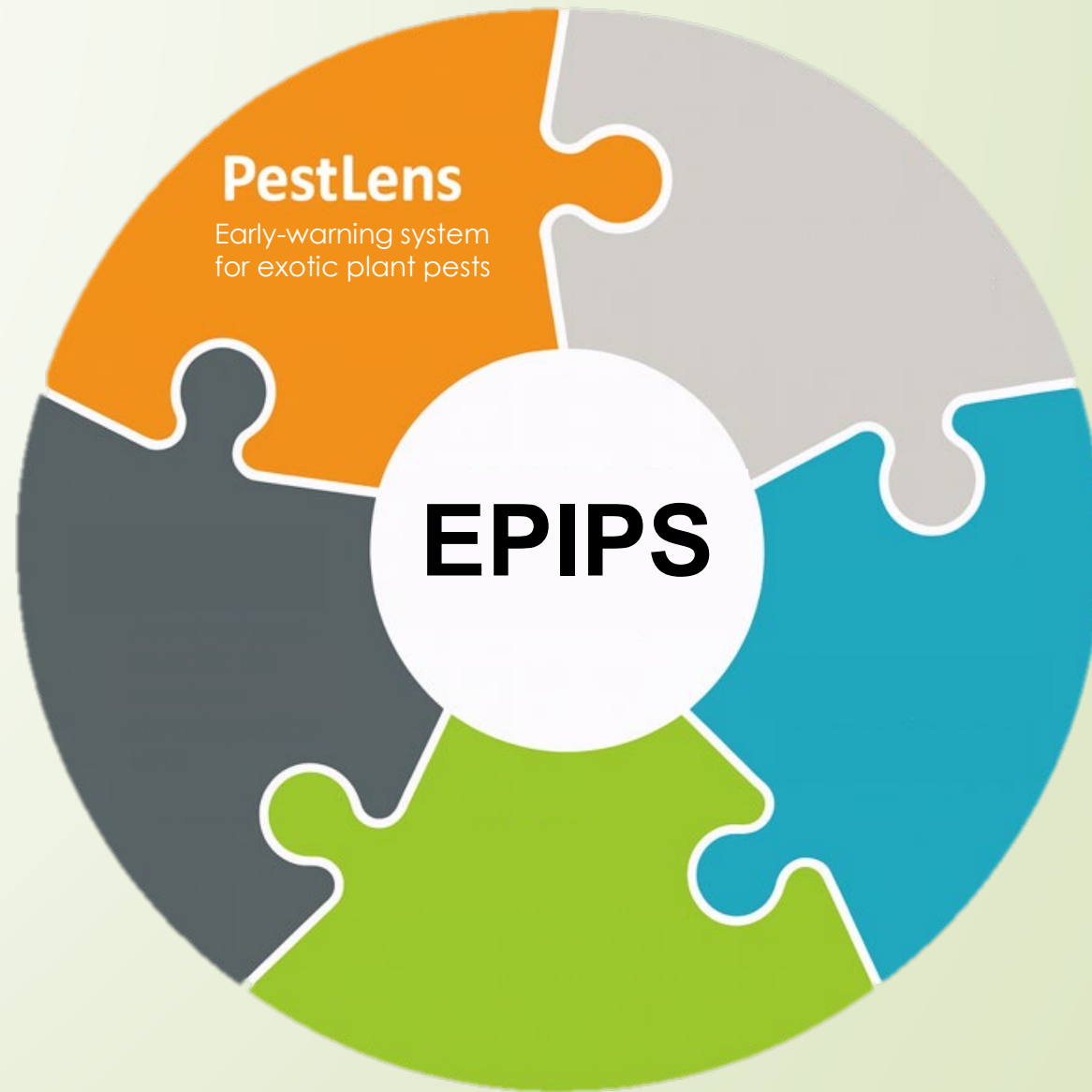


Regulatory Pest Informatics

EPIPS

is a tightly integrated group of systems built and managed by CIPM for USDA Plant Protection and Quarantine (PPQ)

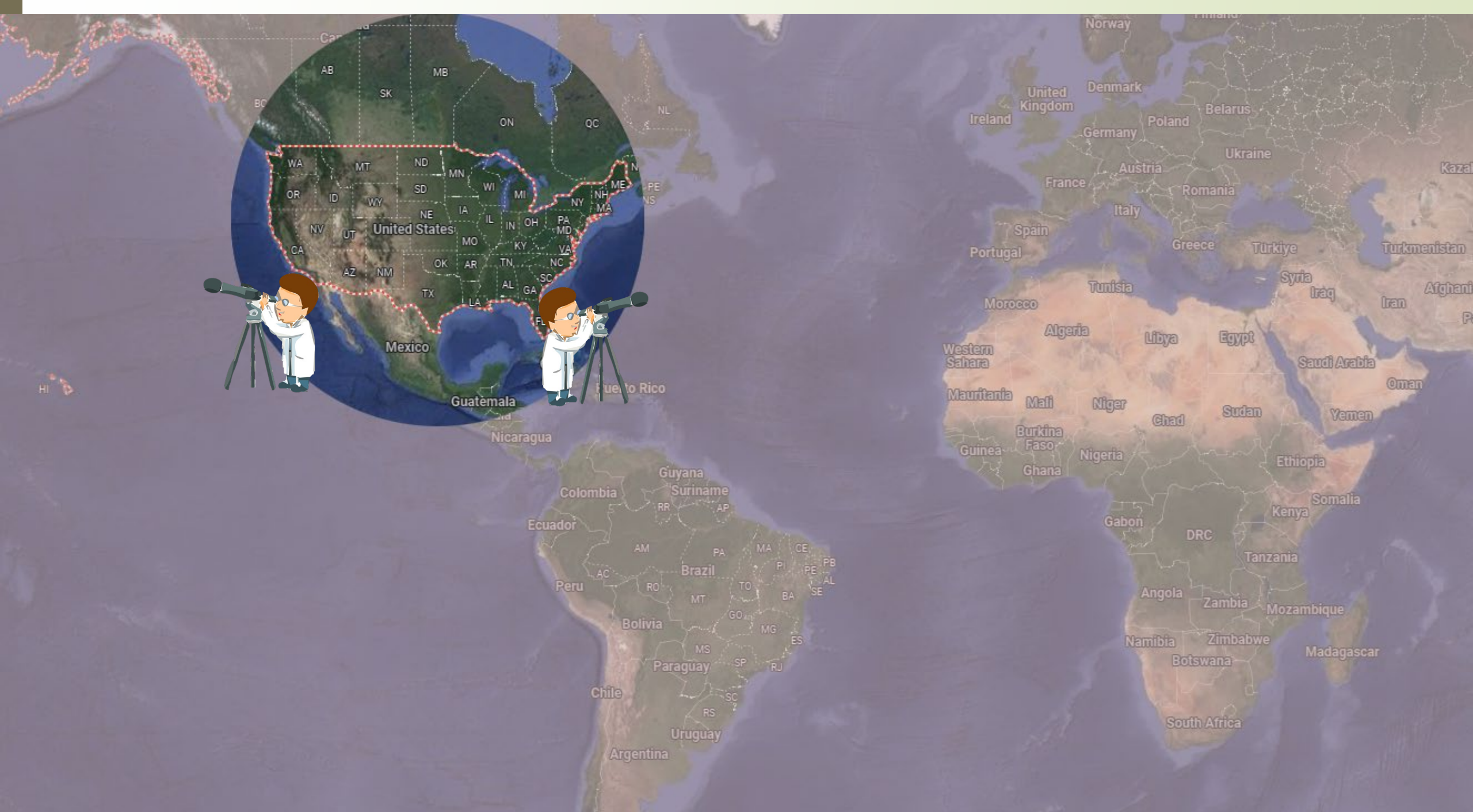






PestLens

PestLens supports our federal stakeholder's efforts to protect U.S. agriculture and natural resources against exotic plant pests by providing an early warning system to notify new pest information to users



Data Collection



A word cloud illustrating various data collection methods. The words are arranged in a circular pattern, each in a different color and orientation. The largest word is 'New Location' in red. Other prominent words include 'New Host' in blue, 'Detection' in orange, 'New Description' in green, 'Research of Interest' in red, 'Naturalization' in teal, 'New Pest' in blue, 'Interception' in purple, 'Eradication' in purple, and 'Other' in red.

Research of Interest

Naturalization

New Pest

New Host

Detection

Other

Eradication

New Description

New Location

Interception

PestLens Article

[View Notification](#)[Save Article](#)[Edit Article](#)

South American palm borer, *Paysandisia archon* (Lepidoptera: Castniidae), established in Switzerland

Reported by PestLens on: Thursday, January 23, 2025

Country: Switzerland

Source: BioInvasions Records

View Count: 104

Event: New Location

During a 2023 survey, South American palm borer, *Paysandisia archon* (Lepidoptera: Castniidae), was found infesting cultivated *Phoenix canariensis* (Canary Island date palm) and *Chamaerops humilis* (dwarf fan palm) plants at multiple locations in Switzerland. Affected plants showed foliage damage, larval tunnels, and larval-feeding sawdust. *Paysandisia archon* was detected in and eradicated from Switzerland from 2010 to 2013.

Paysandisia archon infests economically important palm species, including *Phoenix* spp., *Washingtonia* spp. (fan palm), and *Sabal* spp. (palmetto). *Paysandisia archon* has been reported from other parts of Europe and South America and is not known to occur in the United States.

[View Edits](#)

Reference(s):

1. Fehr, V., A. Minetti, M. Conedera, and G. B. Pezzatti. 2024. First evidence of establishment of the palm borer moth *Paysandisia archon* (Burmeister, 1879) in southern Switzerland. *BioInvasions Records* 13(4):901-908. Last accessed January 23, 2025, from <https://doi.org/10.3391/bir.2024.13.4.05>.

Other resources:

[View more information about *Paysandisia archon* in the Global Pest and Disease Database](#)

Other PestLens articles about this pest:

[South American palm borer, *Paysandisia archon* \(Lepidoptera: Castniidae\), detected in Germany](#)

[First report of the South American palm borer *Paysandisia archon* \(Lepidoptera: Castniidae\) in the Czech Republic](#)

[First report of the South American palm borer *Paysandisia archon* \(Lepidoptera: Castniidae\) in Cyprus and Slovenia](#)



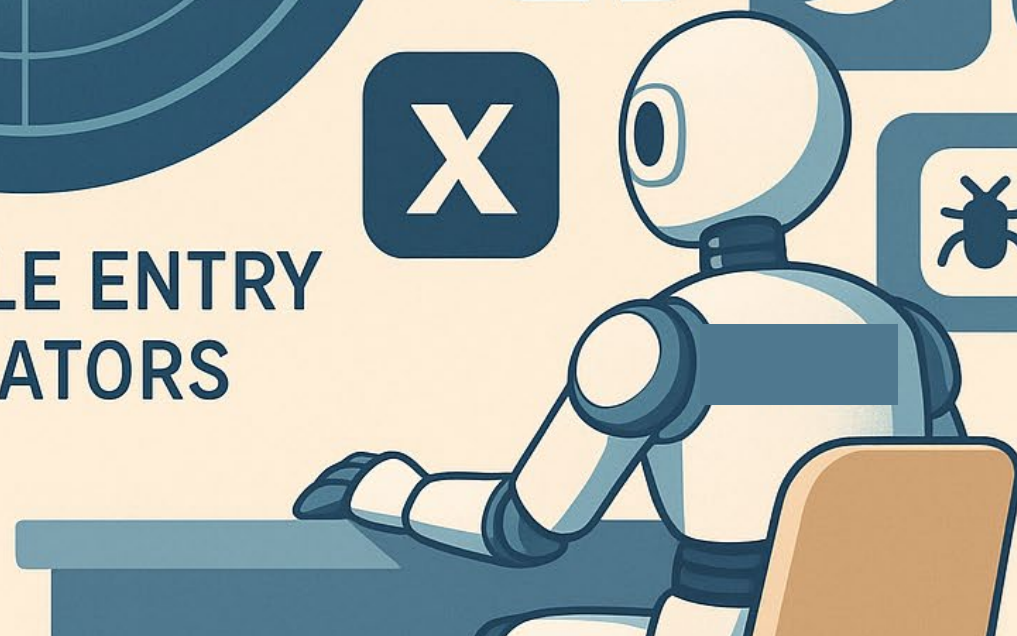
Challenges



**GEOGRAPHIC
PROXIMITY**



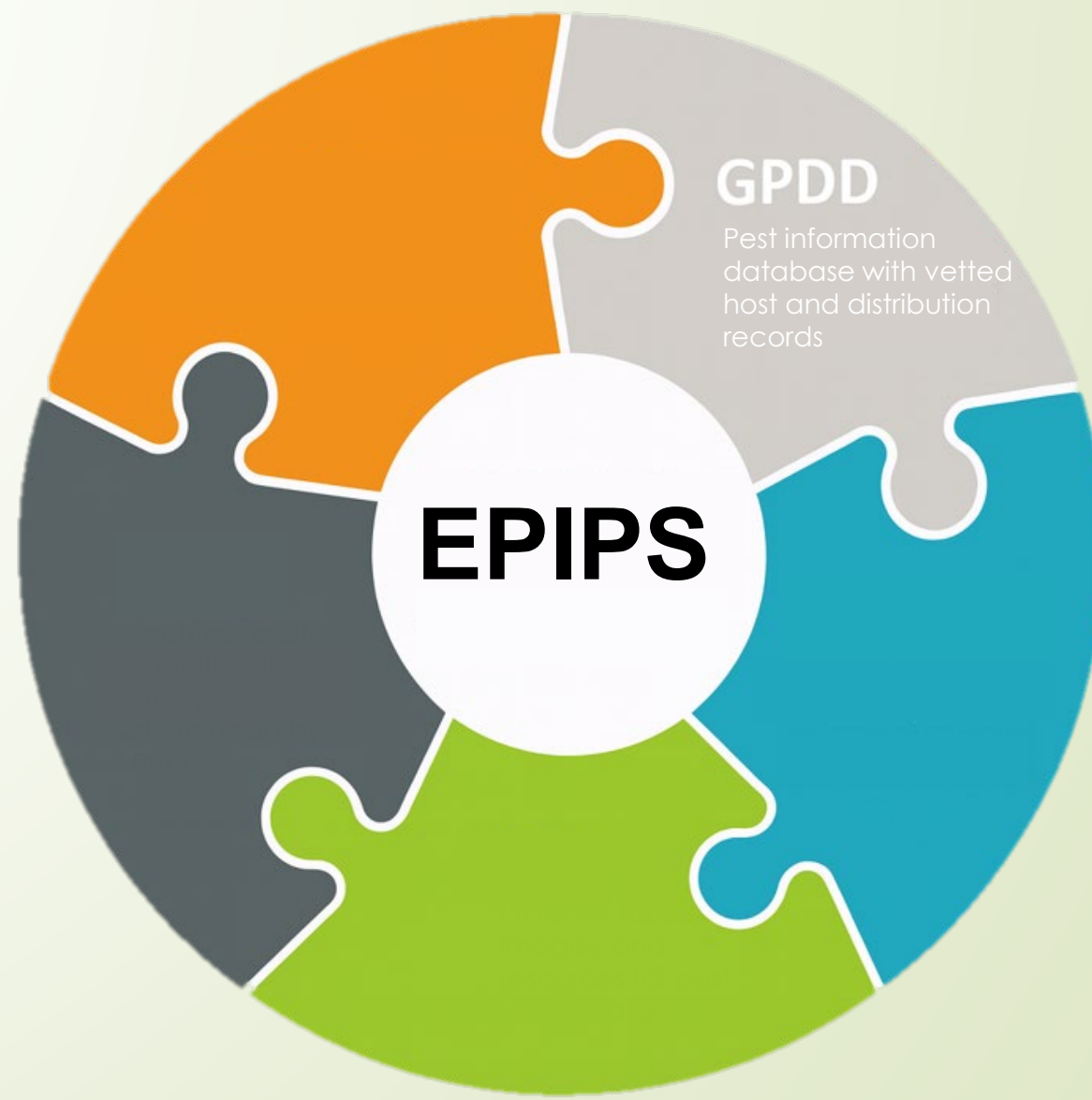
**POSSIBLE ENTRY
INDICATORS**





The illustration shows a person with dark hair and a dark shirt, seen from the back and side, looking at a large computer monitor. The monitor has a dark frame and a light blue screen. On the screen, the text 'EARLY-WARNING SYSTEM NEEDS OPTIMIZATION' is written in bold, white, sans-serif capital letters, arranged in four lines. In the background, there are stylized light blue clouds and a large, light blue triangle with a white exclamation mark inside, suggesting a warning or alert. The overall style is flat and modern.

**EARLY-WARNING
SYSTEM
NEEDS
OPTIMIZATION**





PEST PROFILE



REPORTS

GOVERNMENT
REPORT

PUBLICATIONS

DATASETS

Host Information

Family #	Scientific Name #	Common Name	Evidence of host association is:		
			Direct	Indirect	Erroneous or Disputed
Malvaceae	<i>Abelmoschus esculentus</i> (L.) Moench	Okra	18 , 173	35 , 79 , 279	N/A
Fabaceae	<i>Acacia</i> spp. Mill.	Acacia	393	N/A	N/A
Myrtaceae	<i>Acca sellowiana</i> (O. Berg) Burret	Feijoa	10 , *323 , *18 , *189 , *375	*35 , *344	N/A
Actinidiaceae	<i>Actinidia chinensis</i> Planch. var. <i>chinensis</i>	Actinidia chinensis var. <i>chinensis</i>	121	N/A	N/A
Actinidiaceae	<i>Actinidia deliciosa</i> (A. Chev.) C. F. Liang & A. R. Ferguson	Kiwifruit	121	N/A	N/A

Export to EndNote

Family: Fabaceae

Scientific Name:

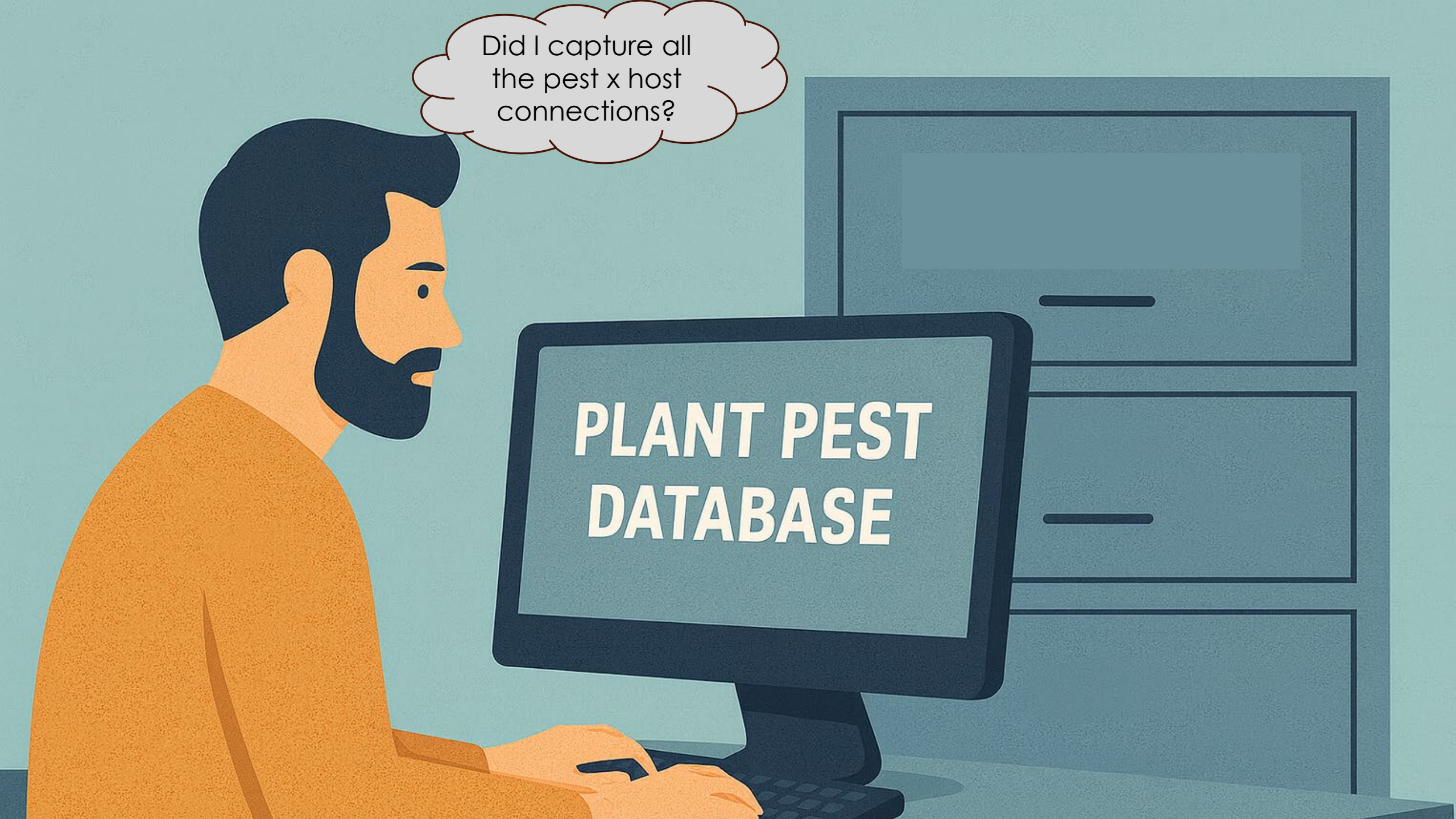
Acacia spp. Mill.

Common Name:

Acacia

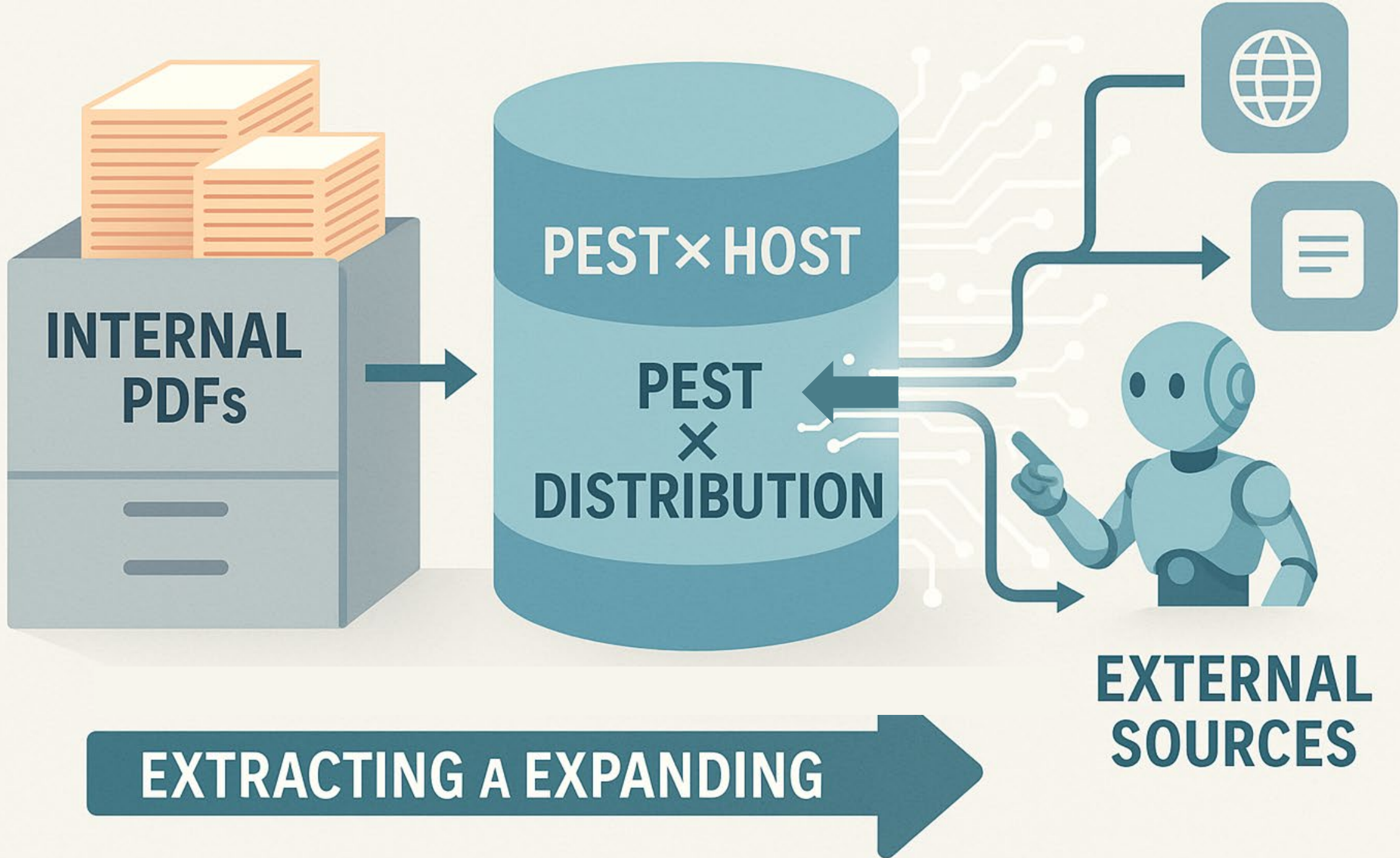
Citation:

(393) (2016, November 10) True Fruit Flies (Diptera, Tephritidae) of the Afrotropical Region: *Bactrocera dorsalis*. Retrieved January 12, 2024 from <http://projects.bebif.be/fruitfly/taxoninfo>.

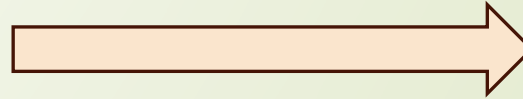


Did I capture all
the pest x host
connections?

PLANT PEST DATABASE



Automated Information Retrieval and Analysis (AIRA)



Host Information

Family #	Scientific Name #	Common Name	Evidence of host association is:		
			Direct	Indirect	Erroneous or Disputed
Malvaceae	<i>Abelmoschus esculentus</i> (L.) Moench	Okra	18 , 173	35 , 79 , 279	N/A
Fabaceae	<i>Acacia</i> spp. Mill.	Acacia	393	N/A	N/A
Myrtaceae	<i>Acca sellowiana</i> (O. Berg) Burret	Feijoa	10 , *323 , *18 , *189 , *375	*35 , *344	N/A
Actinidiaceae	<i>Actinidia chinensis</i> Planch. var. <i>chinensis</i>	Actinidia chinensis var. <i>chinensis</i>	121	N/A	N/A
Actinidiaceae	<i>Actinidia deliciosa</i> (A. Chev.) C. F. Liang & A. R. Ferguson	Kiwifruit	121	N/A	N/A

Export to EndNote

Family: Fabaceae

Scientific Name:

Acacia spp. Mill.

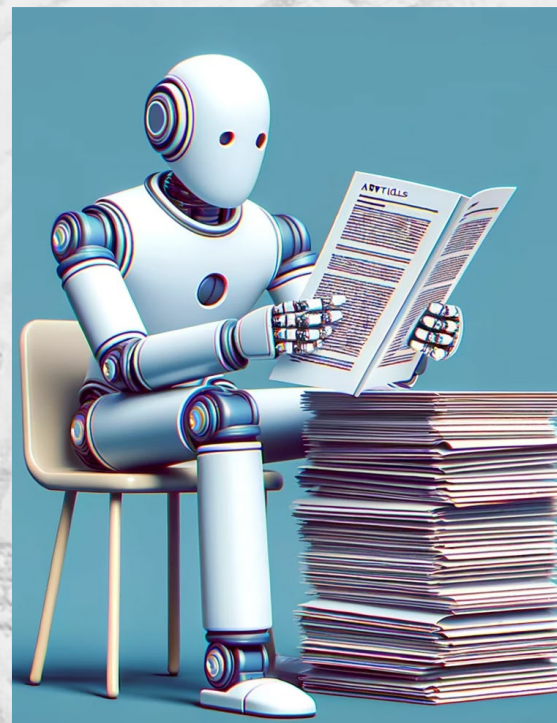
Common Name:

Acacia

Citation:

(393) (2016, November 10) True Fruit Flies (Diptera, Tephritidae) of the Afrotropical Region: *Bactrocera dorsalis*. Retrieved January 12, 2024 from <http://projects.bebif.be/fruitfly/taxoninfo>.

Family 1	Common Name	Scientific Name	Direct	In-Direct	Erroneous	AIRA 
Amaranthaceae	Beet	<i>Beta vulgaris</i>	18,34,67,92			
Amaranthaceae	Spinach	<i>Spinacia oleracea</i>	18,45,73			
Amaranthaceae	Lambsquarters	<i>Chenopodium album</i>	52,56,81,108			
Amaranthaceae	Amaranth, Spiny	<i>Amaranthus spinosus</i>	46,79,113			
Amaranthaceae	Goosefoot, Red	<i>Chenopodium rubrum</i>		18,64		
Amaranthaceae	Good King Henry	<i>Chenopodium bonus-henricus</i>		18		
Asteraceae	Clotbur	<i>Xanthium strumarium</i>	10,38,71,95			
Brassicaceae	Winter Cress	<i>Barbarea vulgaris</i>				A234 ▼ , A267 ▼ , A289 ▼ 
Solanaceae	Black Nightshade	<i>Solanum nigrum</i>				A312 ▼ , A345 ▼ 
Fabaceae	Red Clover	<i>Trifolium pratense</i>				A378 ▼ , A401 ▼ , A423 ▼ , A445 ▼ 
Poaceae	Barnyard Grass	<i>Echinochloa crus-galli</i>				A467 ▼ 
Rosaceae	Wild Strawberry	<i>Fragaria vesca</i>				A489 ▼ , A512 ▼ 



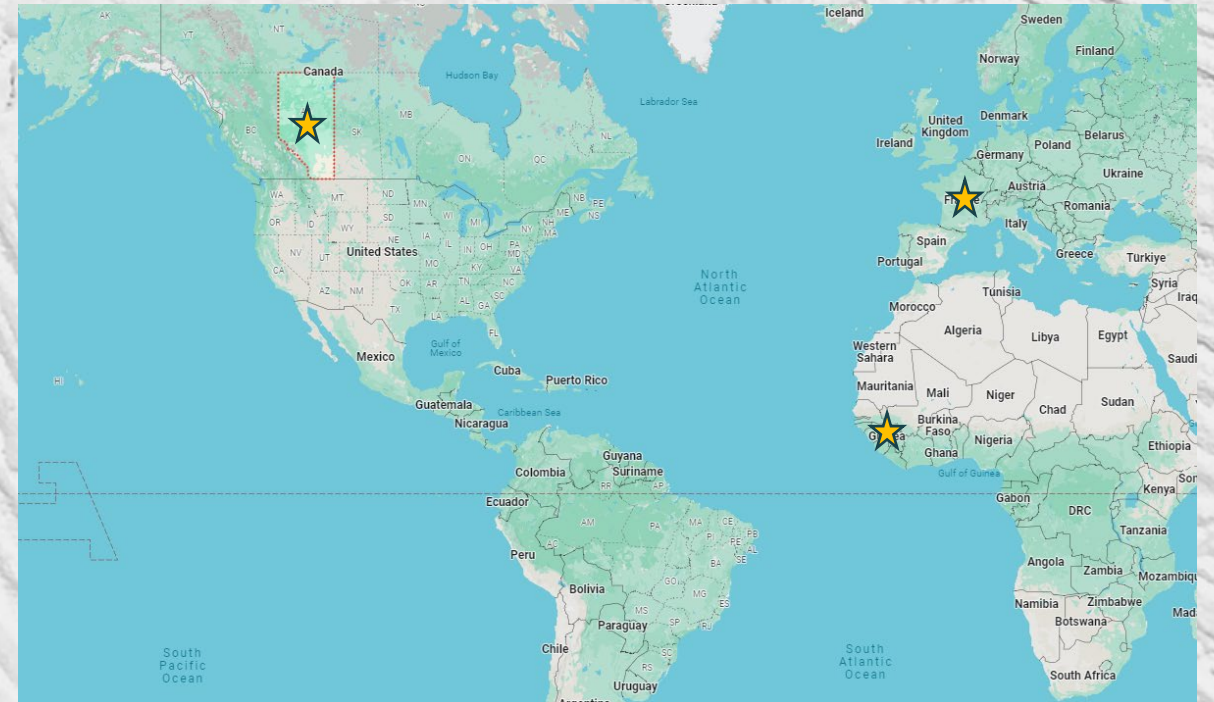
**New hosts
New locations**

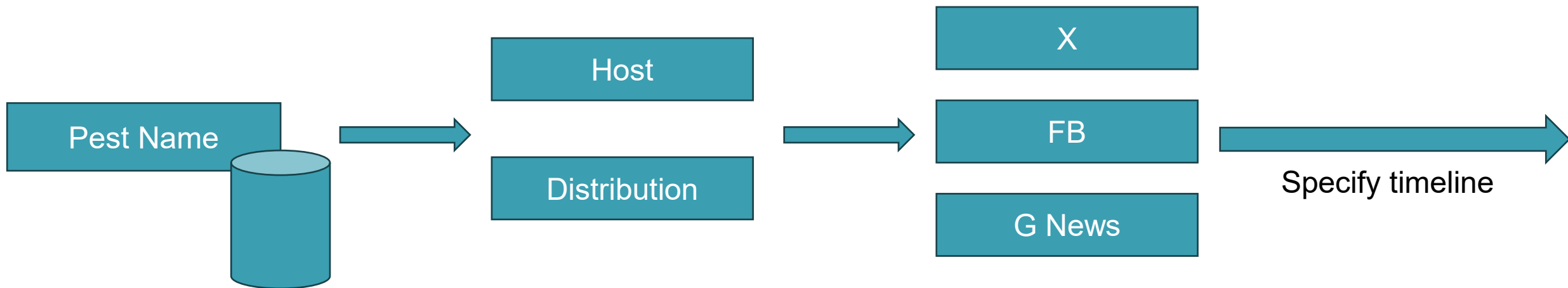
Where are we at?

IPPC pest reports
EPPO reports
EPPO database
NAPPO Pestalerts
Google News
Social media (X, FB)
CAPS survey maps
EDDMaps
iNaturalist
GBIF

LLM deep reports

Google Scholar
Internal PDFs





X

FB

G News



Our criteria





LLM

FACEBOOK	Aceria guerreronis	මයිටාවායාන්තෙක් හඳුන්වන්නේ අසලේ විසිතුරු බෝග වගාවට මෙන්ම සොල් ගස්වල එල දැරීමට බෙහෙවින් බලපාන කෘමියෙකි. * සත්ත්ව වංශයේ ආත් රසෝඩාවෙකු වන මයිටාවා සොල් වගාකරන නිවර්තන කලාපයේ රටවල් බොහොමයක සැතියර සවිති. * සොල් මයිටාවා 1997 වසරේදී සමඟේ රි ලාකාවේ, කල්පිරිය ස් රදේගසෙන් මුලින්ම වාර්තා වී ඇත. මෙම කියවී ඇසට නොසෙකෙන කෘමියා මයික් රෝමීටර 200ක් දිගකින් සහ මයික් රෝමීටර 40ක සමඟේ සමුලකින් යුතු වේ. *** ප් රධාන වගසෙන් සුළගින් සිස් රසෙන් සැතියරයන මෙම හානිය වියළි සහ අතරමැදි කලාපවල දැක ගත හැක. සොල් මයිටාවා, කුඩා සොල් ඇවිරි වල තැටියේ ජීවත් වෙමින්, සොල් ගෙඩිය විශාල වන තෙක්ම එහි යුස පුරාබිම සිදු කරයි. සොල් ඇවිරි එනම් බැව්ට අවදියේ සිට සොල් ගෙඩිය මෝරන අවදිය දක්වා මොවුන් හානි කරන අතර සරාගනය නොවූ සොල් මල් අවදියේ දී හානියක් නොමැත. - වියළි දේශගුණික තත්ත්වය නිසා සොල් ගසේ රසායනික සංයුතියේ ඇති වන වෙනස්වීම් මේ සඳහා සෘජුවම බලපායි. - වියළි කලාපයේ මෙම තත්ත්වයට එක් ර ලෙස මුහුණ දී ඇති අතර, එහි සොල් වගාවෙන් 80%ක් සමඟේ මේ පවුදරට ගොදුරු වී ඇත. - එම නිසා මෙවැනි කාලගුණ තත්ත්වයකදී තෙතමන සංරක්ෂණයේ රම් අනුගමනය කළ යුතුය. මෙම නිසා ගෙඩි කුඩාවීම හා සතුළු වීම වැනි ලක්ෂණ ඇති වේ. මේ හානිය වළක්වා ගැනීමට ගෙන්දුගම් කුඩු මිස් රණය තෝ විලෝපිත මයිටා සැකවි බාවිත කළ හැක. (මේ සිළිබද තවත් තොරතුරු ඔබට මිශ්‍රත සෝස්විට්ට් ඔස්සේ ලබා දීමට බලාසොරොත්තු වන්නෙමු) 	 Sri Lanka
FACEBOOK	Aceria guerreronis	Unknown 11/06/2024 Monitoramento de ataque de Aceria Guerreronis (Ácaro da necrose do coco) empresa Sococo S/A , Moju-pa. 	 Brazil
TWITTER	Aceria guerreronis	@AcarologicalSoc 03/21/2025 11:00:02 Entomopathogenic fungi, Hirsutella thompsonii F.E. Fisher safe mycoacaricide for the management of coconut mite, Aceria guerreronis Keifer in Thailand https://doi.org/10.11158/saa.29.10.1... 	 Thailand
TWITTER	Aceria guerreronis	@Plantacion_50 04/08/2025 08:35:39 Hi. What's about Aceria guerreronis manage and control ? Thanks 	No Country
GNEWS	Aceria guerreronis	Oman Observer Scientific approach essential to tackle coconut mites in Dhofar Coconut trees in Dhofar Governorate are facing significant challenges due to pest infestations. Among these pests, the coconut mite (Aceria.  	 Oman
GNEWS	Aceria guerreronis	Jamaica Gleaner Farmer overjoyed as coconut farm rebounds Lead Stories Farmer Devon Fuller vividly remembers being stumped after three acres of his coconut plantation failed to yield quality coconuts after three...	 Jamaica

FACEBOOK	Aceria guerreronis	මයිටාටා මොසන්තෙන් හඳුන්වන්නේ අසලේ පිස්සුරු බෝග වගාවට මෙන්ම සොල් ගස්වල එලූ දැරීමට බොහෝමත් බලපාන කෘමියෙකි. * සත්ත්ව වගයෙන් ආත් රසෝවාමෙකු වන මයිටාටා සොල් වගාකරන නිවර්තන කලාසයේ රටවල් බොහොමයක සැරිල් සවිතී. * සොල් මයිටාටා 1997 වසරේදී සමස්ත ශ්‍රී ලංකාවේ කල්ලිය ස් රජේගසෙන් මුලින්ම වාර්තා වී ඇත. මෙම පියවි ජූෂට නොසෙසෙන කෘමියා මගින් රෝ මීටර් 200ක් දිගකින් සහ මගින් රෝ මීටර් 40ක සමස්ත පළලකින් යුතු වේ. *** ස් රටාන වගයෙන් සුළුකින් සිස් රසෙන් සැරිලයන මෙම භාතිය වියළි සහ අතරමැදි කලාසවල දැන ගත හැක. සොල් මයිටාටා, කුඩා සොල් ඇවිලි වල හැරීමේ ජීවත් වෙමින්, සොල් ගෙඩිය වගාලා වන තෙක්ම එහි යුෂ උරාලීම සිදු කරයි. සොල් ඇවිලි එනම් බැවිලි අවිදීමේ සිට සොල් ගෙඩිය මෝරන අවදිය දක්වා මොවුන් භාතී කරන අතර සරාගනය නොවූ සොල් මල් අවිදීමේ දී භාතියක් නොමැතිය. - වියළි රජේගසක හත්ත්වය නිසා සොල් ගසේ රසායනික සංයුතියේ ඇති වන වෙනස්වීම් මේ සඳහා සෘජුවම බලපායි. - වියළි කලාසයේ මෙම හත්ත්වයට උග් ර ලෙස මුහුණ දී ඇති අතර, එහි සොල් වගාවෙන් 80%ක් සමස්ත මේ උවදුරට ගොදුරු වී ඇත. - එම නිසා මෙවැනි කාලගුණ හත්ත්වයකදී තෙතමන සංරක්ෂණයක් රම් අනුගමනය කළ යුතුය. මෙම නිසා ගෙඩි කුඩාවීම හා පලදු වීම වැනි ලක්ෂණ ඇති වේ. මේ භාතිය වළක්වා ගැනීමට ගෙන්දගම් කුඩු මේ රණය හෝ විලෝපිත මයිටා සැකවි බාවිත කළ හැක. (මේ සිද්ධිබදු හවත් තොරතුරු ඔබට මිශ්‍රත සෝස්විටුව ඔස්සේ ලබා දීමට බලාපොරොත්තු වන්නෙමු) 	 Sri Lanka
FACEBOOK	Aceria guerreronis	Unknown 11/06/2024 Monitoramento de ataque de Aceria Guerreronis (Ácaro da necrose do coco) empresa Sococo S/A, Moju-pa. 	 Brazil
TWITTER	Aceria guerreronis	@AcarologicalSoc 03/21/2025 11:00:02 Entomopathogenic fungi, Hirsutella thompsonii F.E. Fisher safe mycoacaricide for the management of coconut mite, Aceria guerreronis Keifer in Thailand https://doi.org/10.11158/saa.29.10.1... 	 Thailand
TWITTER	Aceria guerreronis	@Plantacion_50 04/08/2025 08:35:39 Hi. What's about Aceria guerreronis manage and control ? Thanks 	No Country
GNEWS	Aceria guerreronis	Oman Observer Scientific approach essential to tackle coconut mites in Dhofar Coconut trees in Dhofar Governorate are facing significant challenges due to pest infestations. Among these pests, the coconut mite (Aceria). 	 Oman
GNEWS	Aceria guerreronis	Jamaica Gleaner Farmer overjoyed as coconut farm rebounds Lead Stories Farmer Devon Fuller vividly remembers being stumped after three acres of his coconut plantation failed to yield quality coconuts after three...	 Jamaica



Is some of this info already present in GPDD?

Just show me the rest



Where are we at?

IPPC pest reports

EPPO reports

EPPO database

NAPPO Pestalerts

Google News

Social media (X, FB)

CAPS survey maps

EDDMaps

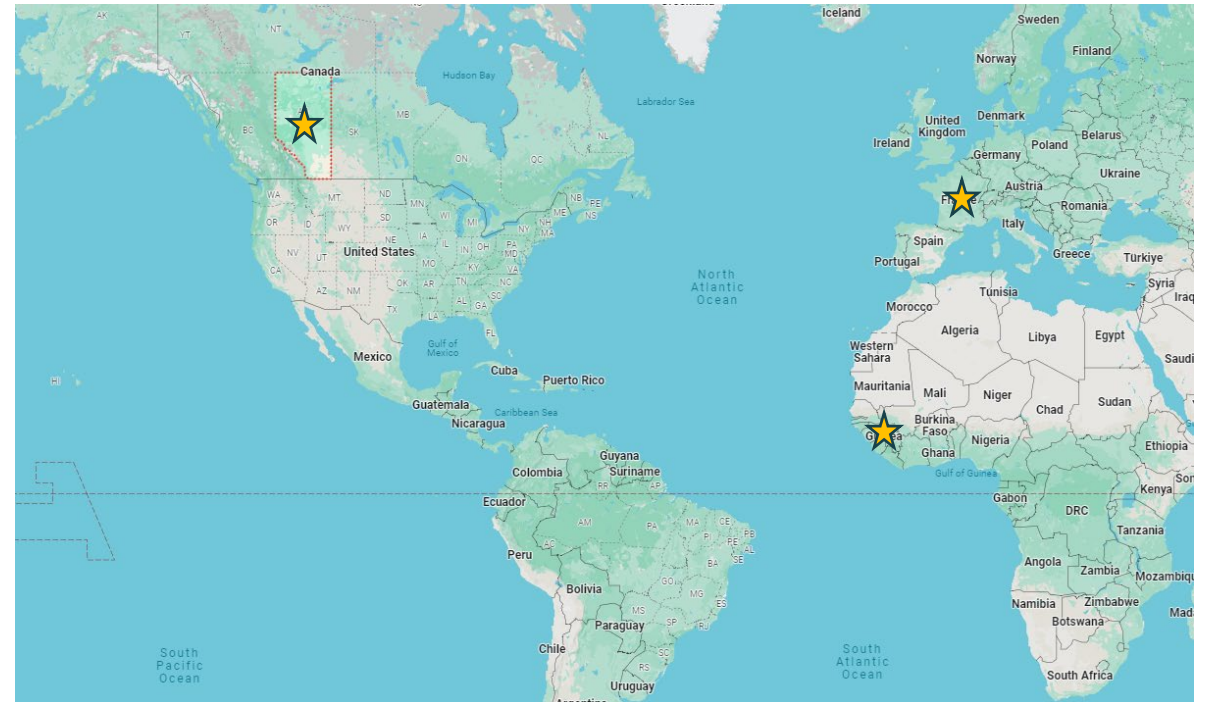
iNaturalist

GBIF

LLM deep reports

Google Scholar

Internal PDFs

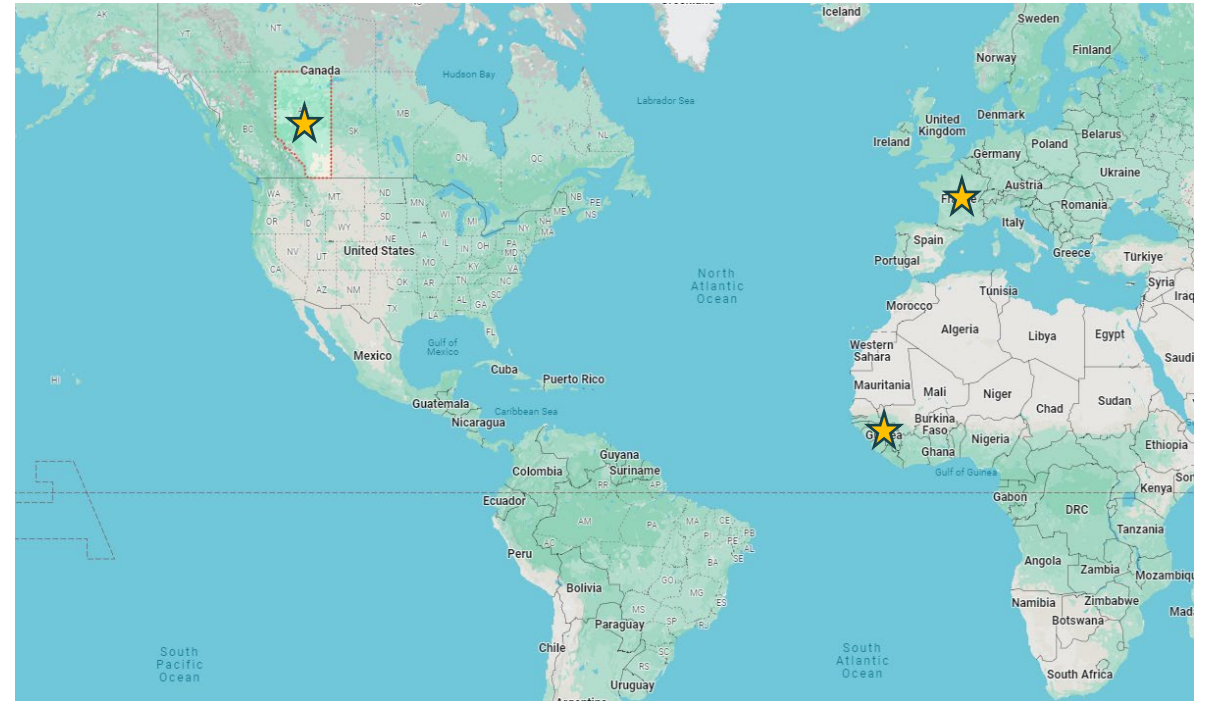


Where are we at?

IPPC pest reports
EPPO reports
EPPO database
NAPPO Pestalerts
Google News
Social media (X, FB)
CAPS survey maps
EDDMaps
iNaturalist
GBIF

LLM deep reports

Google Scholar
Internal PDFs

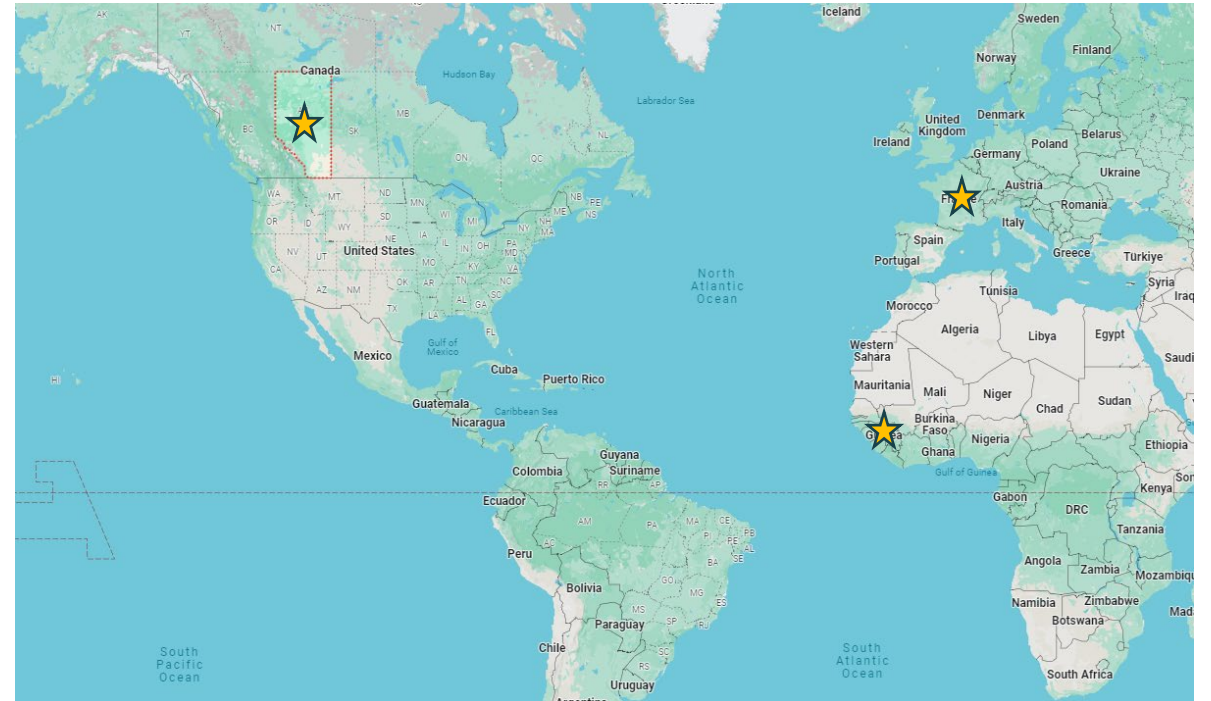


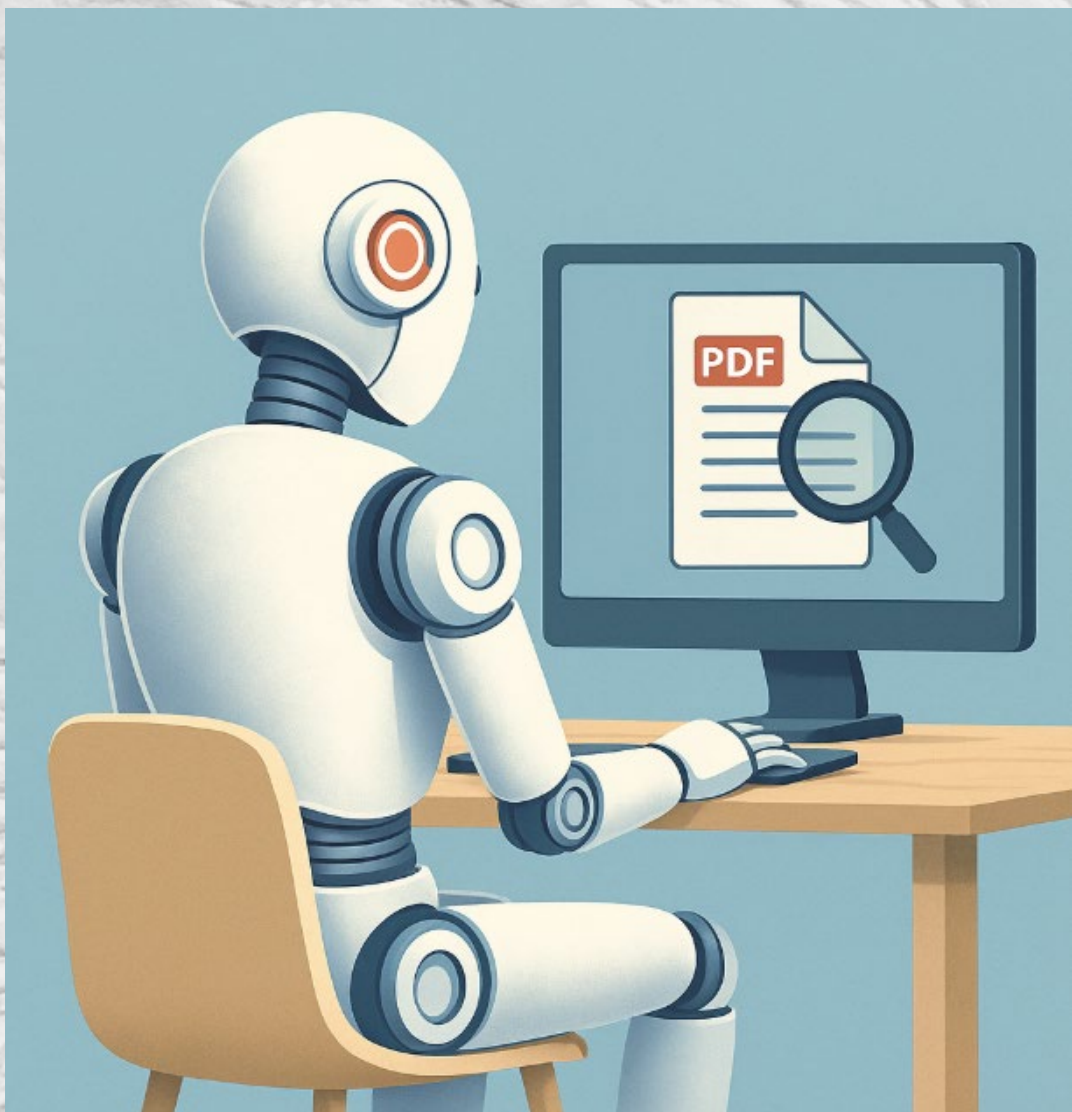
Where are we at?

IPPC pest reports
EPPO reports
EPPO database
NAPPO Pestalerts
Google News
Social media (X, FB)
CAPS survey maps
EDDMaps
iNaturalist
GBIF

LLM deep reports

Google Scholar
Internal PDFs





ChatGPT



Claude



Gemini



Government Reports
PRAs
Other Internal PDFs

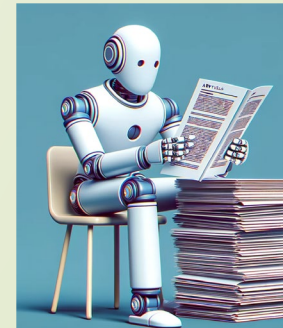
Extract
Text

Prompts

Local



LLM





LLM-Powered Agricultural Intelligence System

CIPM and PPQ

System Trigger: Pest Information Update Request

When a user requests pest information to be updated, our system automatically begins data collection from multiple sources

Fall Armyworm (*Spodoptera frugiperda*)

 Update Pest Info

 Run Complete Demo

 Reset System

 System Analytics

1 Multi-Source Data Extraction

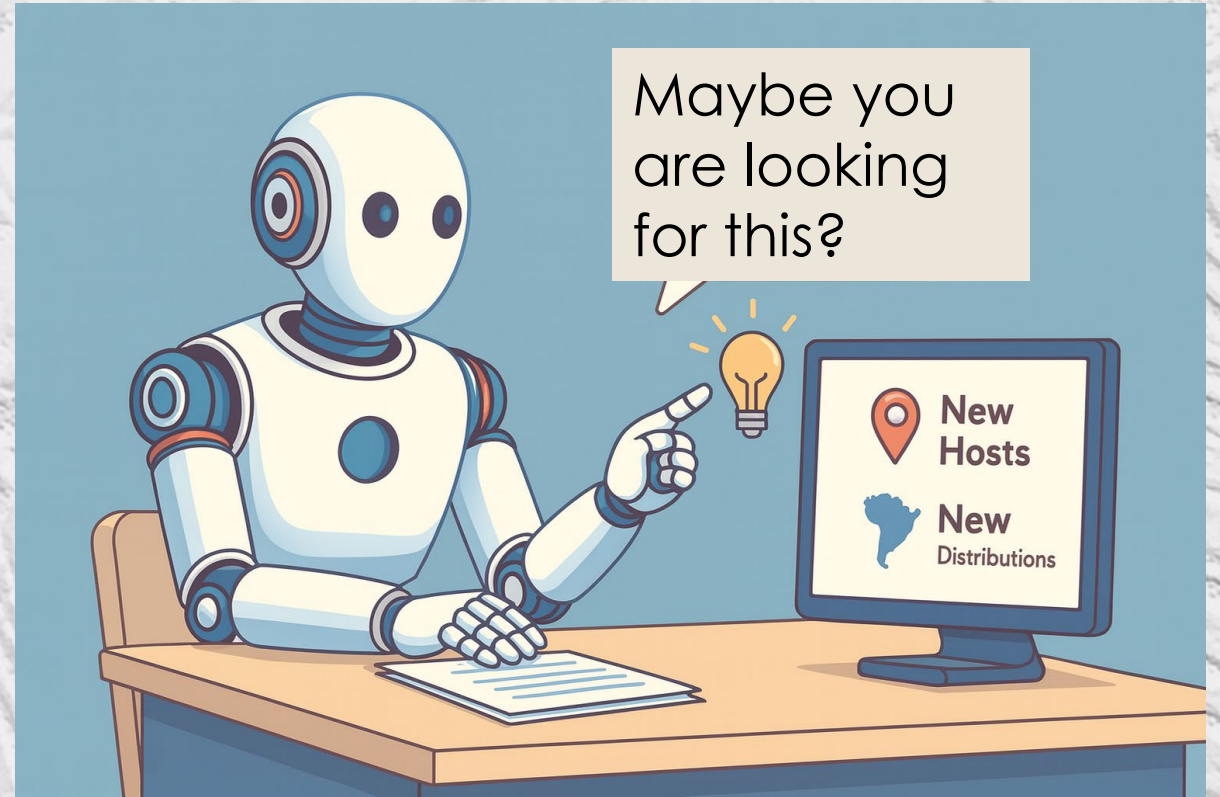
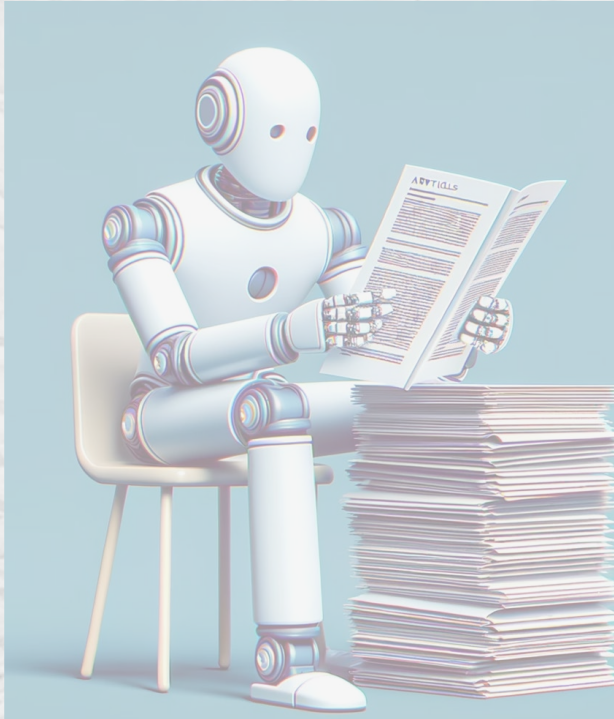
Ready

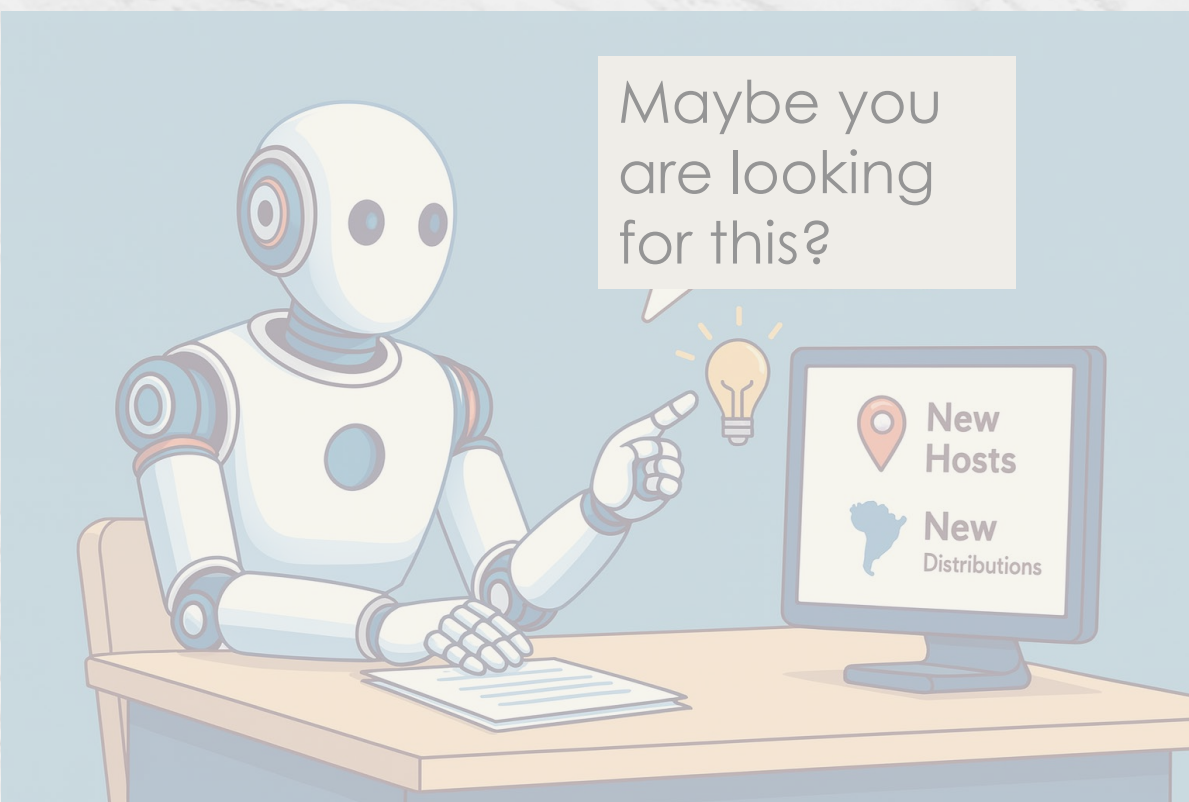
2 MongoDB Data Storage

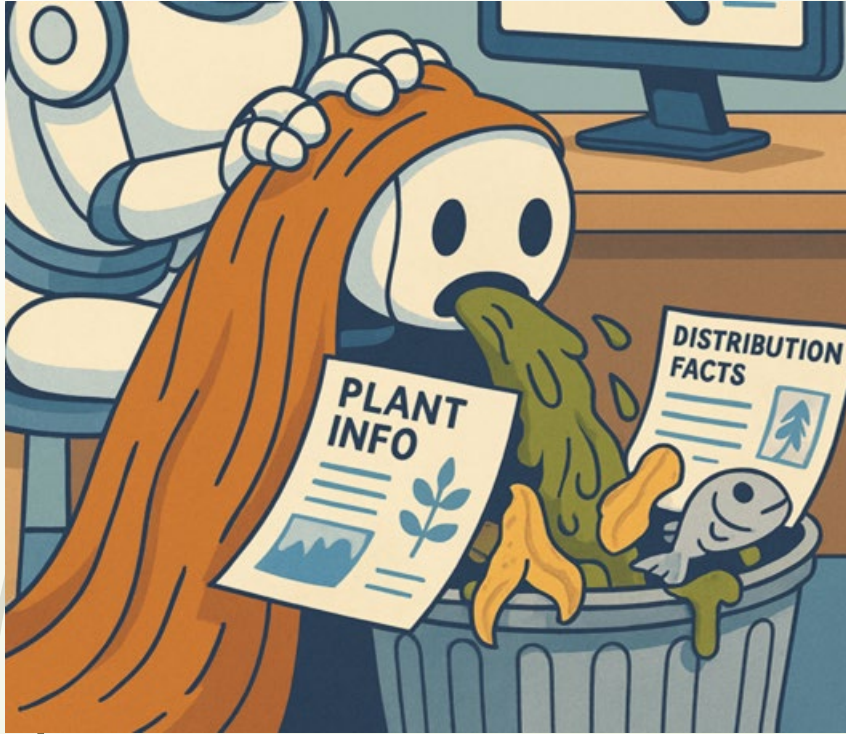
Ready

3 Semantic Search & Data Retrieval

Ready









Government Reports
PRAs
Other Internal PDFs

Extract
Text

OCR
Format

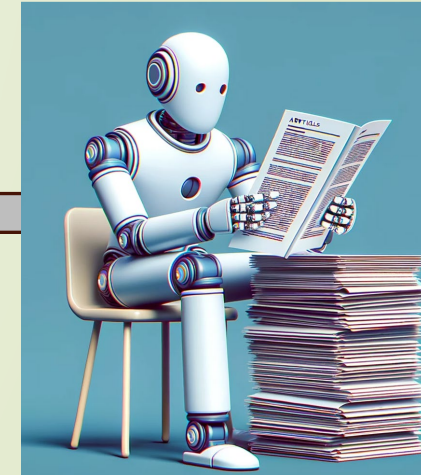
Prompts

Local



LLM

Chunking strategy





Temperature Playground

Explore how creativity settings impact AI response quality

By: Godshen Robert

Creativity Level (Temperature)



Medium Creativity (0.5)



Balanced Approach



AIRA Vector RAG Architecture



Data Preparation Pipeline

A

Government Documents Source

AIRA focuses on government agricultural documents as the primary source for pest intelligence - providing authoritative, verified information from official agencies.



Government Documents



Official Reports



Agency Publications



Research Studies

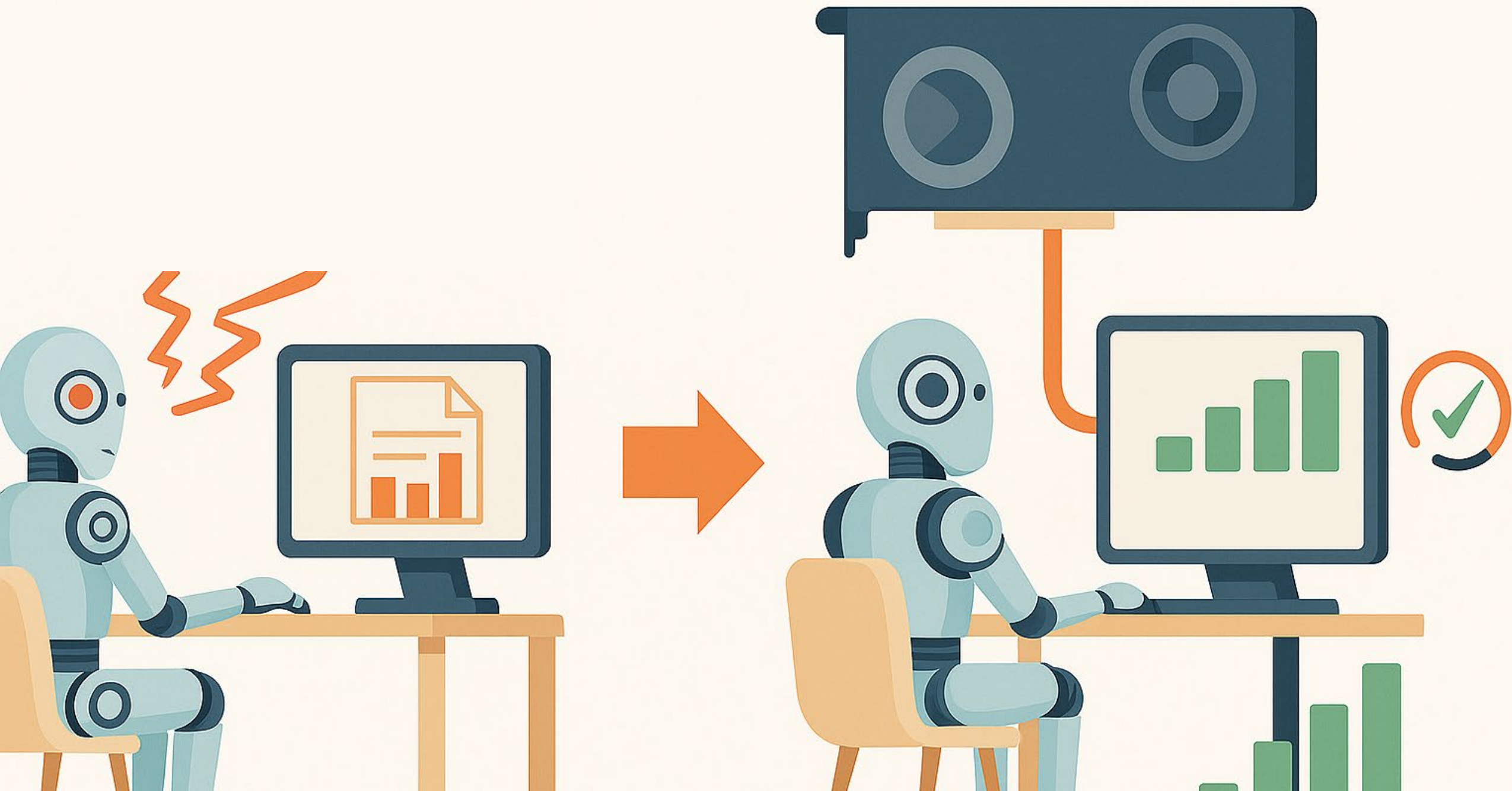
Focus Source: Government agricultural documents from agencies like USDA, state extension services, and international agricultural organizations

B

Split Documents into Chunks

Government documents are split into manageable chunks because LLM models have context limits. Each chunk preserves meaningful context while staying within processing boundaries.

Why Split into Chunks? • LLM models have context limits (e.g., 4K, 8K, or 32K tokens) • Large documents exceed these limits • Chunks preserve semantic meaning in smaller pieces **Example Chunk:** "Fall armyworm (*Spodoptera frugiperda*) damage has been confirmed on rice paddies in northern Thailand. Local extension agents in Chiang Mai province report larvae feeding on *Oryza sativa* with 15-20% yield losses recorded in March 2024." **Metadata:** {source: "government_doc", location: "thailand", date: "2024-03"}





Data Sources

Extraction process

Pre-processing

Prompts & LLM

Outputs & Processing

Final Output



Can we
automate it?

EVALUATE PIPELINES

Pipeline settings

Configure New Pipeline

PDF Properties

PDF Type

Modern OCRed

PDF Access Level

Public

Pre-processing

Library

Dots.ocr

LLM Input Format

text + images

LLM Configuration

LLM Model

Gemma 3 - 27B

Prompt Template

Prompt 1

Analyze the following document to identify pest-host and pest-

Pipeline Name

A descriptive name for this configuration

Cancel

Save to Library

PROMPT

PDF TYPE

LIBRARY

Prompt 1

Scanned

Dots.ocr

Prompt 2

Modern OCRed

torchvision

Filter by PDF Type

All

SOURCE

PDF TYPE

Peer-reviewed

Scanned

Peer-reviewed

Modern OCRed

Government Report

Modern OCRed

Government Report

Modern OCRed

Pipeline Performance examples

Overall Pipeline Performance

Aggregated performance metrics for each pipeline across all experiments.

PIPELINE ↕	F1 SCORE ▼	PRECISION ↕	RECALL ↕	HALLUCINATION ↕
Gemma3_Scanned	0.915	0.940	0.891	0.082
Phi4_Modern_OCRed	0.880	0.851	0.912	0.115

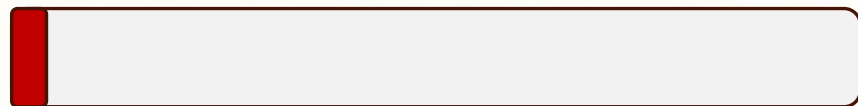


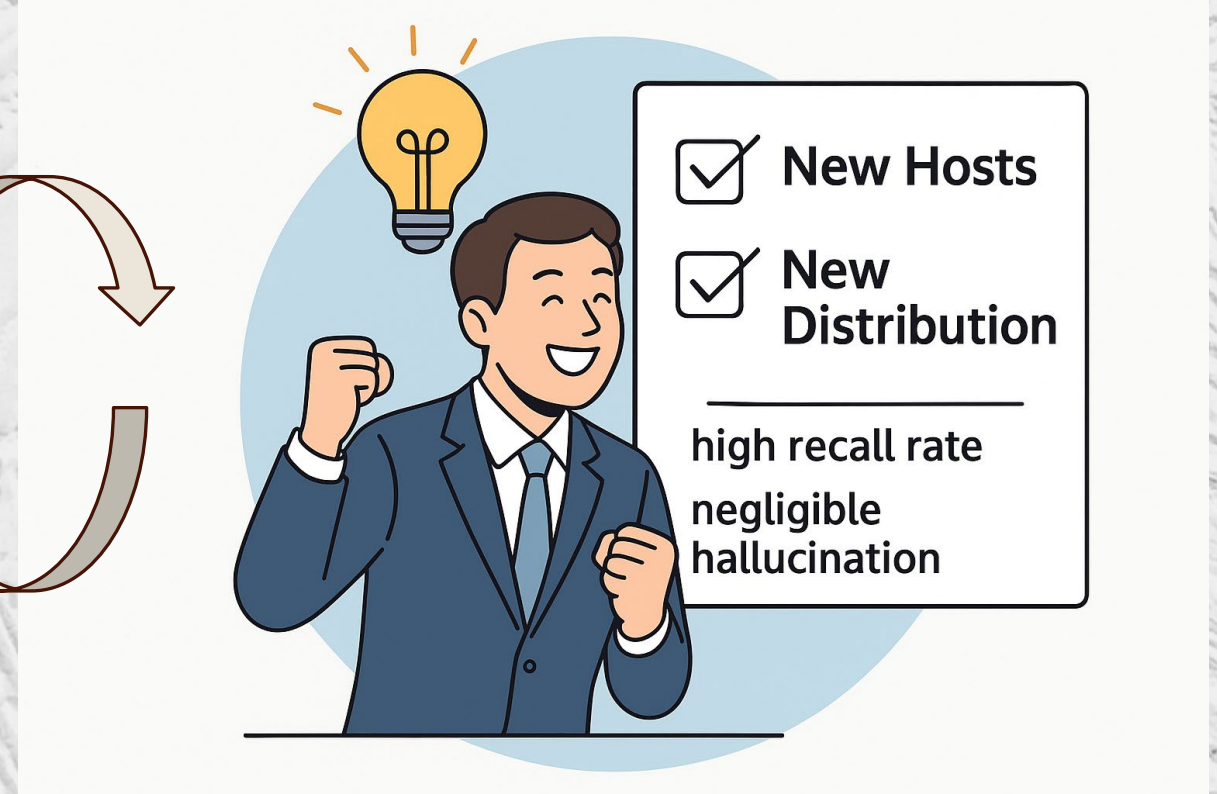
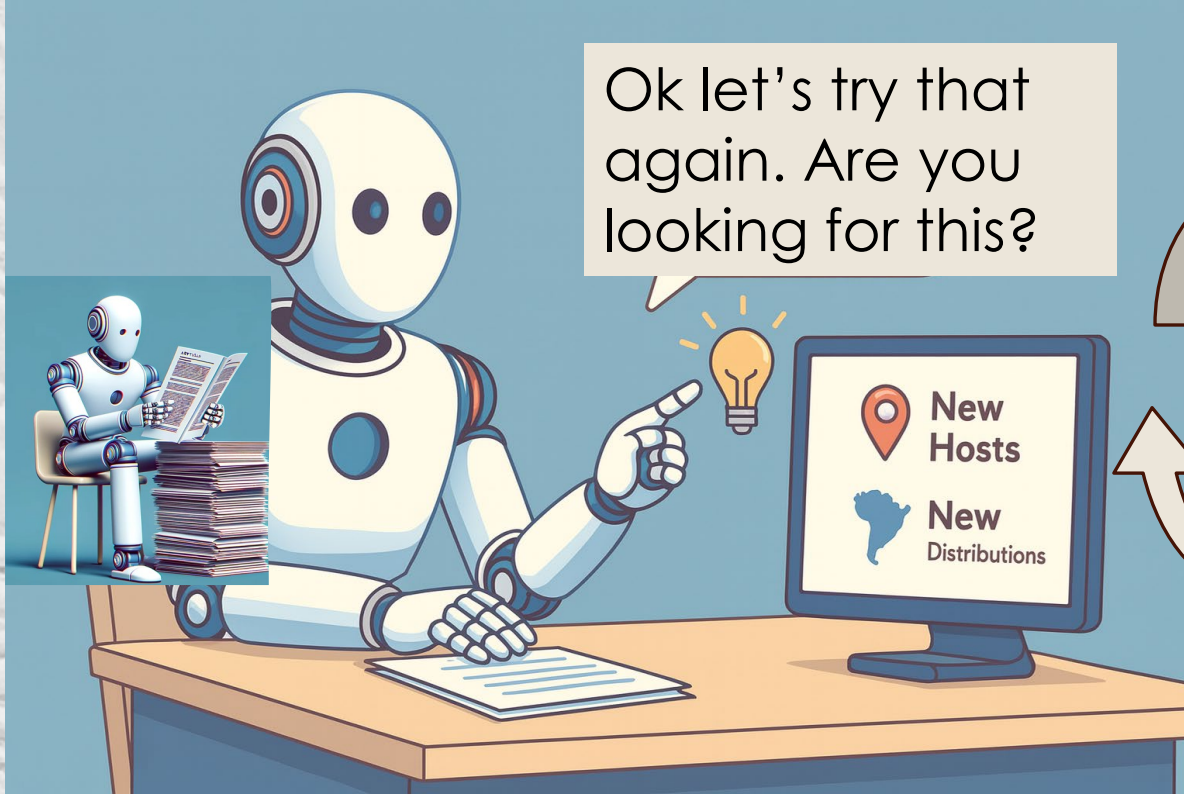
EVALUATE PIPELINES

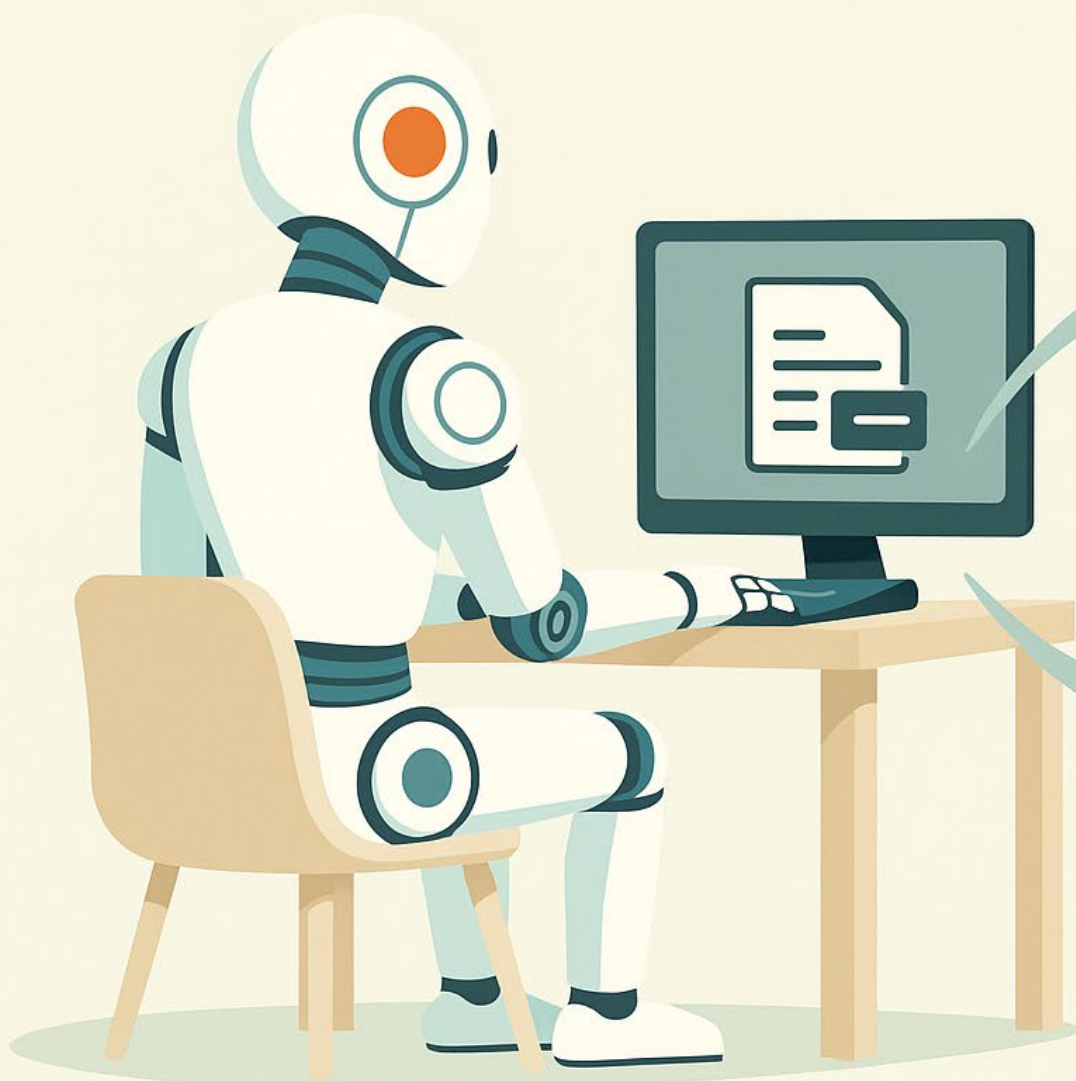
ACCURACY



HALLUCINATIONS







**Early Warning
(PestLens)**



**Pest Information
(GPDD)**



Expected AIRA output integration into Pest Database

GPDD Host Information Table

Interactive simulation showing how AIRA-discovered host associations would appear in the GPDD table. The AIRA column contains references to literature discovered through the RAG pipeline.

AIRA Human-in-the-Loop Workflow:

- **Click any AIRA reference** to review and categorize the evidence
- **Accept as Direct:** Move validated reference to Direct evidence column
- **Accept as Indirect:** Move to Indirect evidence if circumstantial
- **Mark as Erroneous:** Flag incorrect associations for exclusion
- Database analysts can batch process multiple AIRA discoveries efficiently

Family T	Common Name	Scientific Name	Direct	In-Direct	Erroneous	AIRA 📄
Amaranthaceae	Beet	<i>Beta vulgaris</i>	18,34,67,92			
Amaranthaceae	Spinach	<i>Spinacia oleracea</i>	18,45,73			
Amaranthaceae	Lambsquarters	<i>Chenopodium album</i>	52,56,81,108			
Amaranthaceae	Amaranth, Spiny	<i>Amaranthus spinosus</i>	46,79,113			
Amaranthaceae	Goosefoot, Red	<i>Chenopodium rubrum</i>		18,64		
Amaranthaceae	Good King Henry	<i>Chenopodium bonus-henricus</i>		18		
Asteraceae	Clotbur	<i>Xanthium strumarium</i>	10,38,71,95			
Brassicaceae	Winter Cress	<i>Barbarea vulgaris</i>				

A234 ▼, A267 ▼, A289 ▼ 🔗



NC STATE UNIVERSITY

Center for Integrated Pest Management

