FCAvizIR: Exploring Relational Data Set's Implications Using Metrics and Topics

Lola Musslin, Alexandre Bazin, Marianne Huchard, Pierre Martin, Pascal Poncelet, Vincent Raveneau, Arnaud Sallaberry

FCAvizIR: Exploring Relational Data Set's Implications using Metrics and Topics

Lola Musslin, Alexandre Bazin, Marianne Huchard, Pierre Martin,
Pascal Poncelet, Vincent Raveneau, Arnaud Sallaberry

LIRMM AIDA AMIS @ Univ. Montpellier, CNRS, CIRAD, Univ. Paul-Valéry Montpellier 3

CONCEPTS'24

Table of Contents

Context

Analyzing data through implications in FCA/RCA setting

Context

Analyzing data through implications in FCA/RCA setting

- “What can we remember and gather from the data forgetting the objects? Answers to this question would help in abstracting the meaning of things” [Duquenne, 2013]

Context

Analyzing data through implications in FCA/RCA setting

- “What can we remember and gather from the data forgetting the objects? Answers to this question would help in abstracting the meaning of things” [Duquenne, 2013]
- Multi-relational datasets (Relational Concept Analysis)

Context

Analyzing data through implications in FCA/RCA setting

- “What can we remember and gather from the data forgetting the objects? Answers to this question would help in abstracting the meaning of things” [Duquenne, 2013]
- Multi-relational datasets (Relational Concept Analysis)

A Few Current Applications

Context

Analyzing data through implications in FCA/RCA setting

- “What can we remember and gather from the data forgetting the objects? Answers to this question would help in abstracting the meaning of things” [Duquenne, 2013]
- Multi-relational datasets (Relational Concept Analysis)

A Few Current Applications

- Environmental domain (Knomana): Understanding relations between therapeutic plants, pests (bioagressors), and affected organisms to find or hypothesize solutions in phytotherapy [Mahrach et al., 2021]

Context

Analyzing data through implications in FCA/RCA setting

- “What can we remember and gather from the data forgetting the objects? Answers to this question would help in abstracting the meaning of things” [Duquenne, 2013]
- Multi-relational datasets (Relational Concept Analysis)

A Few Current Applications

- Environmental domain (Knomana): Understanding relations between therapeutic plants, pests (bioagressors), and affected organisms to find or hypothesize solutions in phytotherapy [Mahrach et al., 2021]
- Software Engineering domain: Extracting Variability in Software Product Lines, in a relational context [Carbonnel et al., 2019]

Context

Analyzing data through implications in FCA/RCA setting

- “What can we remember and gather from the data forgetting the objects? Answers to this question would help in abstracting the meaning of things” [Duquenne, 2013]
- Multi-relational datasets (Relational Concept Analysis)

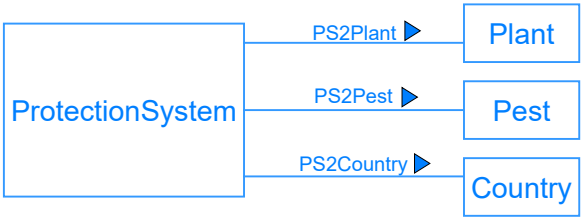
A Few Current Applications

- Environmental domain (Knomana): Understanding relations between therapeutic plants, pests (bioagressors), and affected organisms to find or hypothesize solutions in phytotherapy [Mahrach et al., 2021]
- Software Engineering domain: Extracting Variability in Software Product Lines, in a relational context [Carbonnel et al., 2019]
- Digital Debate domain: Stimulating debates and enriching the lexico-semantic KB [Ben Sassi et al., 2024]

Table of Contents

RCA in a nutshell

Dataset model: Several formal contexts (4) and relational contexts (3)

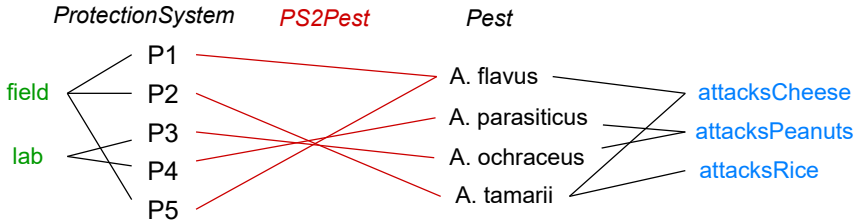


Excerpt from Knomana [Silvie et al., 2021]

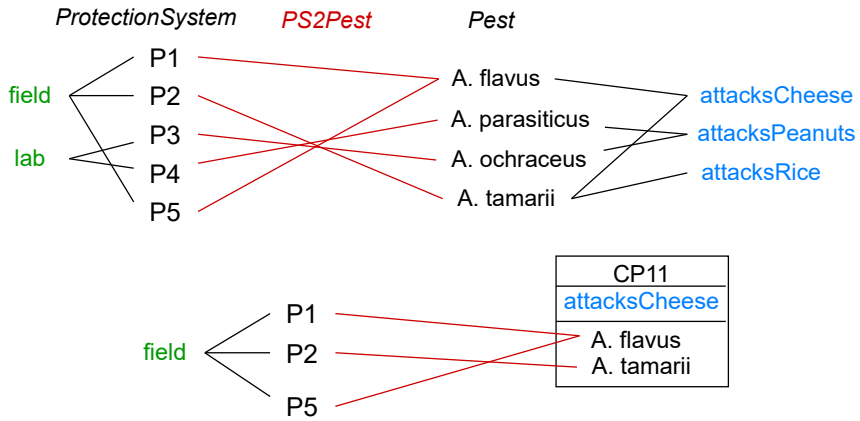
P1 (Protection System in the field) developed in Namibia (Country): Cinnamomum zeylanicum (Plant) is used to control Aspergillus flavus (Pest)

A protection system indicates
a plant usage to control a pest (bioagressor) in a given country

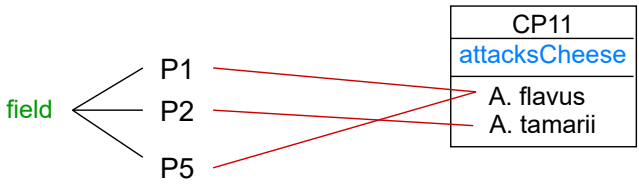
From the dataset to Pest concept CP11



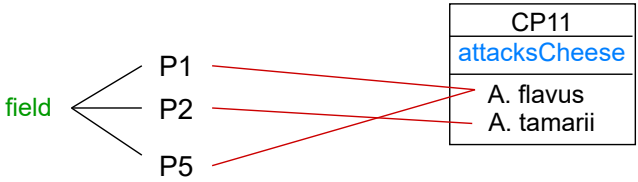
From the dataset to Pest concept CP11



From Pest concept CP11 to Relational Attribute ∃PS2Pest(CP11) and Protection System concept CPS2



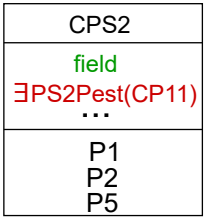
From Pest concept CP11 to Relational Attribute ∃PS2Pest(CP11) and Protection System concept CPS2

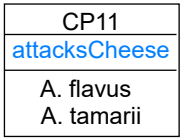


CPS2
field ∃PS2Pest(CP11) ...
P1 P2 P5

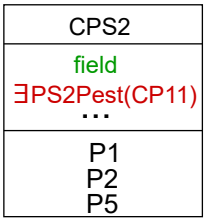
CP11
attacksCheese
A. flavus A. tamarii

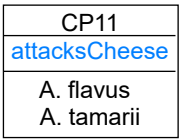
Simplification of Relational Attribute ∃PS2Pest(CP11) using simplified intent attacksCheese

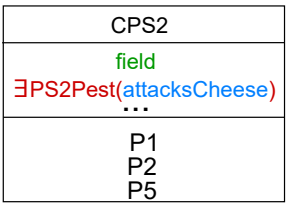


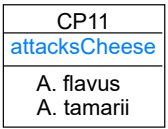


Simplification of Relational Attribute ∃PS2Pest(CP11) using simplified intent attacksCheese









Application to the whole model (4 formal contexts and 3 relational contexts): extended formal context Protection System

ProtSyst	field	lab	∃PS2Pest(attacksRice)	∃PS2Pest(attacksCheese)	∃PS2Pest(attacksPeanuts)	∃PS2Pest(aspergillus)	∃PS2Country(australAfrica)	∃PS2Country(westernAfrica)	∃PS2Plant(contraceptive)	∃PS2Plant(antidysenteric)	∃PS2Plant(lauraceae*)	∃PS2Plant(asteraceae**)	∃PS2Plant(aromatic&evergreen)
P1	x			x		x	x				x		x
P2	x		x	x		x	x				x		x
P3		x			x	x		x				x	x
P4		x			x	x		x	x			x	x
P5	x			x		x	x			x		x	x

Implications are computed from the extended formal context Protection System

A few implications of the Duquenne-Guigues base of implications [Guigues and Duquenne, 1986]:

∃PS2Pest(aspergillus),∃PS2Plant(lauraceae&comestible&applicOil),∃PS2Plant(aromatic&evergreen) ⇒ field,∃PS2Pest(attacksCheese),∃PS2Country(australAfrica)

∃PS2Pest(aspergillus),∃PS2Country(westernAfrica),∃PS2Plant(aromatic&evergreen) ⇒ lab,∃PS2Pest(attacksPeanuts),∃PS2Plant(asteraceae&toxic&applicEssentialOil&applicExtract)

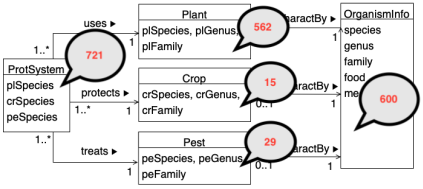
∃PS2Pest(attacksPeanuts),∃PS2Pest(aspergillus),∃PS2Plant(aromatic&evergreen) ⇒ lab,∃PS2Country(westernAfrica),∃PS2Plant(asteraceae&toxic&applicEssentialOil&applicExtract)

(...)

Table of Contents

Facing the number and complexity of implications

Example on a real data model [Mahrach et al., 2021]



Duquenne-Guigues bases of implications (# implications with support $\neq 0$)

- 1391 on protection systems
- 1815 on plants
- 70 on crops
- 80 on pests
- 1168 on organism information

Facing the number and complexity of implications

Principles

- Visualization to support result exploration

Facing the number and complexity of implications

Principles

- Visualization to support result exploration
- Juxtapose and coordinate different views [Baldonado et al., 2000, Roberts, 2007]

Facing the number and complexity of implications

Principles

- Visualization to support result exploration
- Juxtapose and coordinate different views [Baldonado et al., 2000, Roberts, 2007]
- “overview first, zoom and filter, then details on demand” [Shneiderman, 1996]

Facing the number and complexity of implications

Principles

- Visualization to support result exploration
- Juxtapose and coordinate different views [Baldonado et al., 2000, Roberts, 2007]
- “overview first, zoom and filter, then details on demand” [Shneiderman, 1996]
- Separate relations and their targets in relational attributes

Facing the number and complexity of implications

Principles

- Visualization to support result exploration
- Juxtapose and coordinate different views [Baldonado et al., 2000, Roberts, 2007]
- “overview first, zoom and filter, then details on demand” [Shneiderman, 1996]
- Separate relations and their targets in relational attributes
- Use the content of the implications

Requirements

- **R0** Filtering with various metrics to rapidly localize implications with certain values of these metrics

Requirements

- **R0** Filtering with various metrics to rapidly localize implications with certain values of these metrics
- **R1** Clustering based on common elements (relations or attributes) in their premise or conclusion.

Requirements

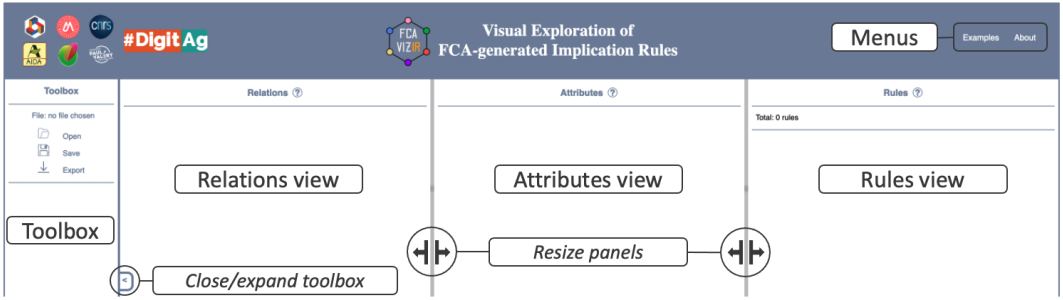
- **R0** Filtering with various metrics to rapidly localize implications with certain values of these metrics
- **R1** Clustering based on common elements (relations or attributes) in their premise or conclusion.
- **R2** Estimating the quantity of implications in the formed groups.

Requirements

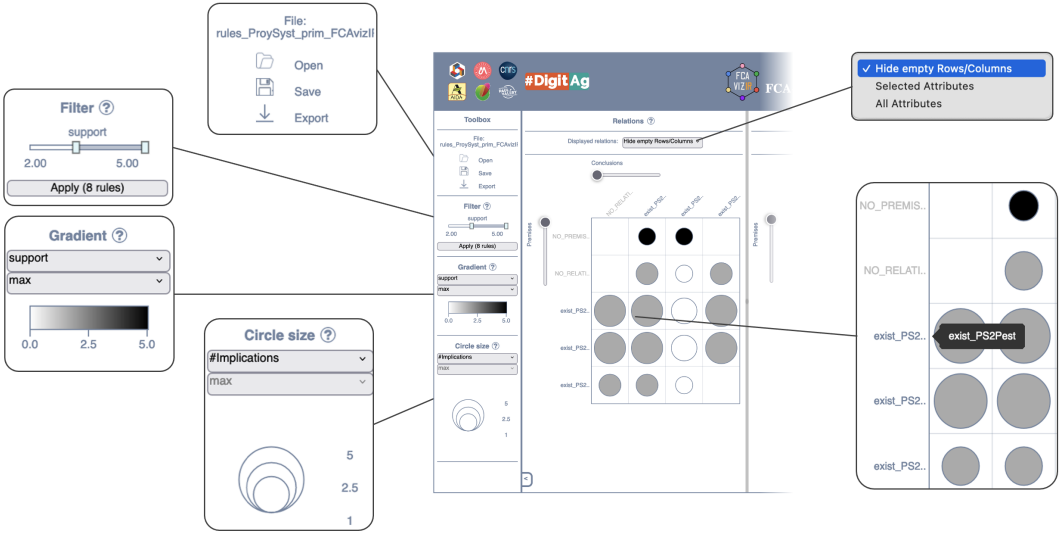
- **R0** Filtering with various metrics to rapidly localize implications with certain values of these metrics
- **R1** Clustering based on common elements (relations or attributes) in their premise or conclusion.
- **R2** Estimating the quantity of implications in the formed groups.
- **R3** Navigating among topic-based implication groups by refining/enlarging the selection step-by-step.

Table of Contents

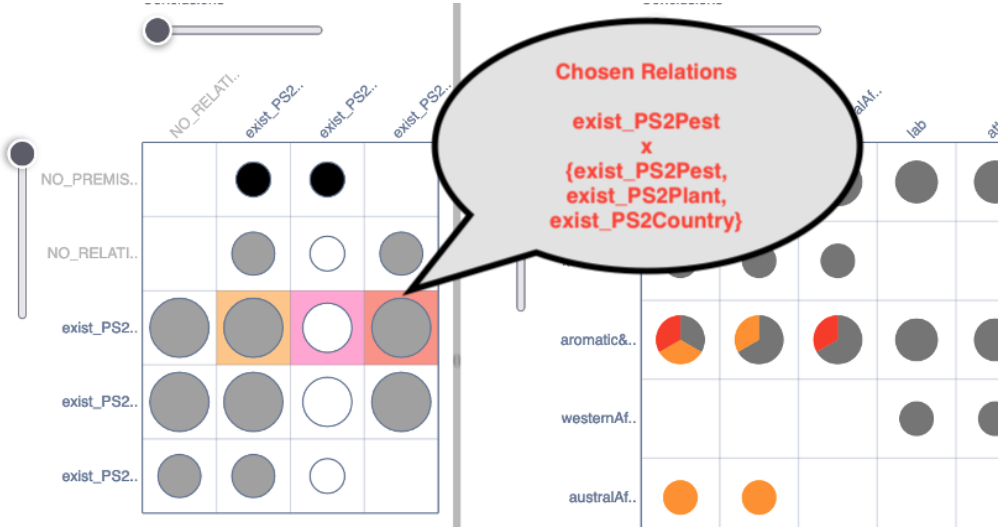
Overview


#DigitAg
Visual Exploration of
FCA-generated Implication Rules
Menus
Examples
About
Toolbox
File: no file chosen
Open
Save
Export
Relations
Attributes
Rules
Total: 0 rules
Relations view
Attributes view
Rules view
Toolbox
Resize panels
Close/expand toolbox

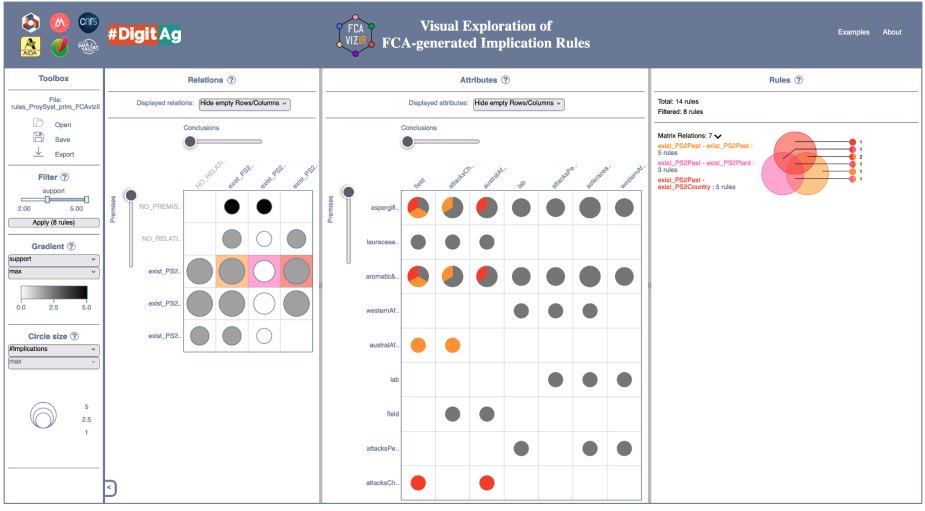
Toolbox / Parameters



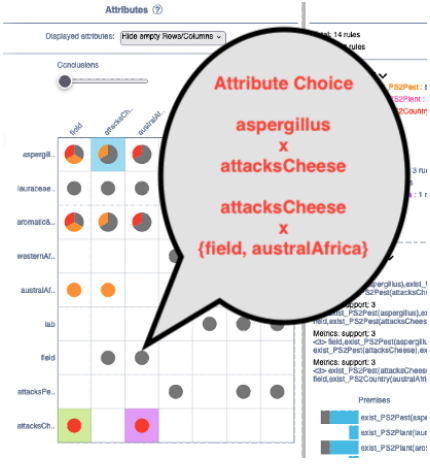
Relation Selection



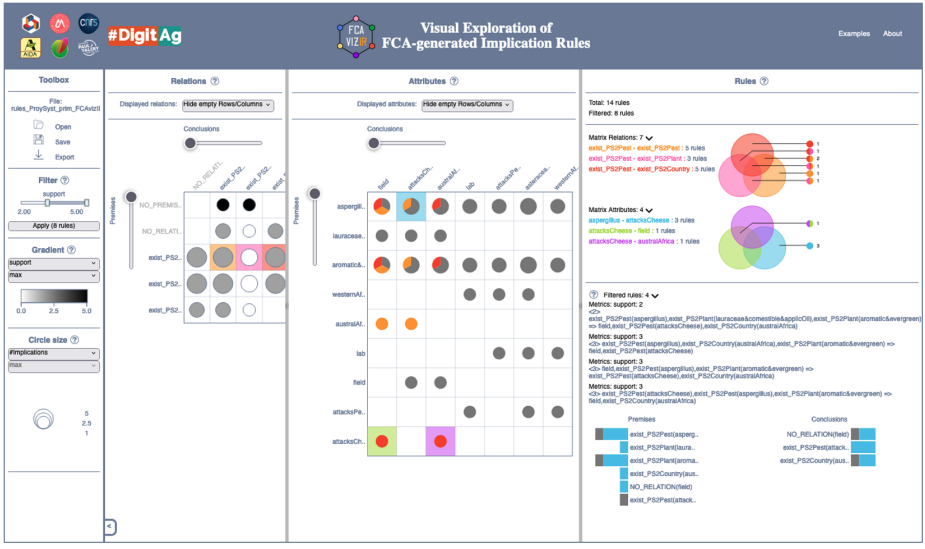
Relation Selection (result)



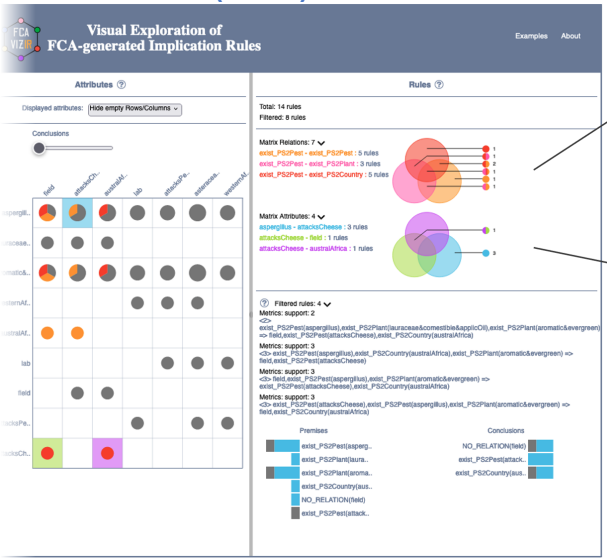
Attribute Selection



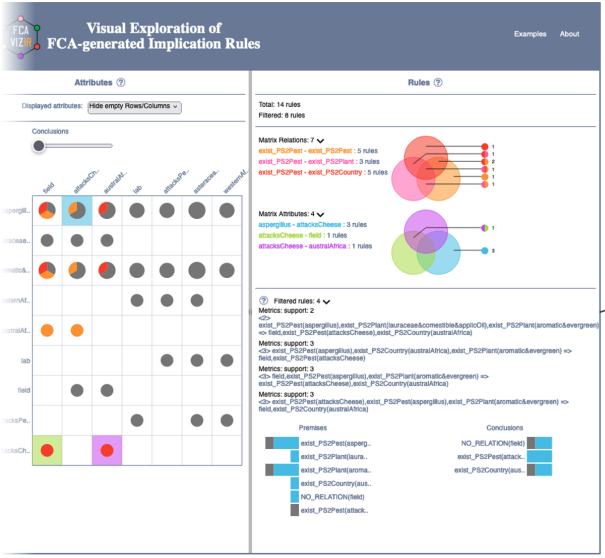
Attribute Selection (result)

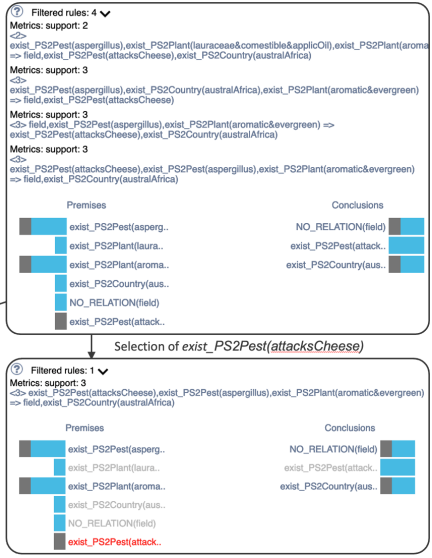


Attribute Selection (detail)



Premise Selection





Premise Selection (result)

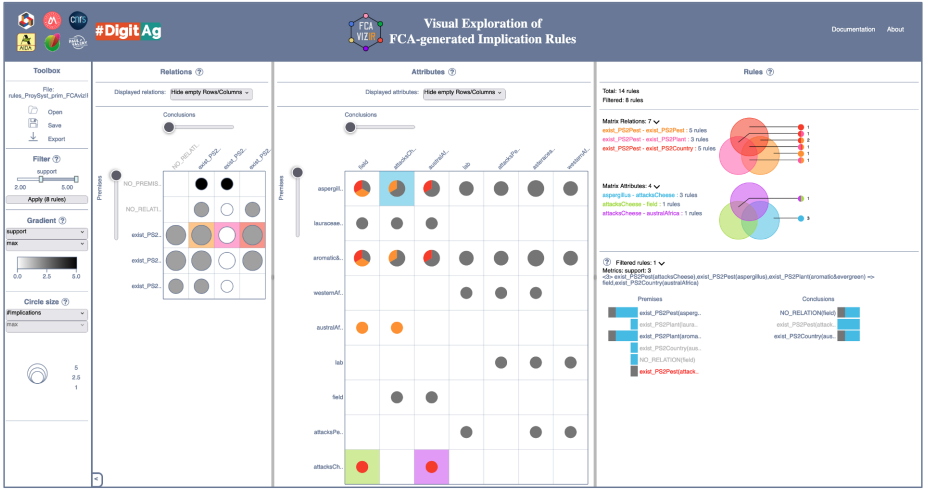


Table of Contents

Conclusion

FCAvizIR

- Visual approach to assist exploration of RCA implications

Conclusion

FCAvizIR

- Visual approach to assist exploration of RCA implications
- Web platform

Conclusion

FCAvizIR

- Visual approach to assist exploration of RCA implications
- Web platform
- Interactive views, filtering (R0), group by shared items (R1), group size visualization (R2), topic-based navigation (R3)

Future work

Implications and models specificities

Future work

Implications and models specificities

- Focus on particular cases: cyclic models, models with long paths

Future work

Implications and models specificities

- Focus on particular cases: cyclic models, models with long paths
- Extension to Triadic and Triadic RCA [Bazin et al., 2024]

Future work

Implications and models specificities

- Focus on particular cases: cyclic models, models with long paths
- Extension to Triadic and Triadic RCA [Bazin et al., 2024]
- Taxonomies on attributes

Future work

Implications and models specificities

- Focus on particular cases: cyclic models, models with long paths
- Extension to Triadic and Triadic RCA [Bazin et al., 2024]
- Taxonomies on attributes

Connect to other tools FCA4J (`https://www.lirmm.fr/fca4j/`), RCAviz (`https://rcaviz.lirmm.fr/`), etc.

Future work

Implications and models specificities

- Focus on particular cases: cyclic models, models with long paths
- Extension to Triadic and Triadic RCA [Bazin et al., 2024]
- Taxonomies on attributes

Connect to other tools FCA4J (`https://www.lirmm.fr/fca4j/`), RCAviz (`https://rcaviz.lirmm.fr/`), etc.

Future work

Implications and models specificities

- Focus on particular cases: cyclic models, models with long paths
- Extension to Triadic and Triadic RCA [Bazin et al., 2024]
- Taxonomies on attributes

Connect to other tools FCA4J (`https://www.lirmm.fr/fca4j/`), RCAviz (`https://rcaviz.lirmm.fr/`), etc.

Future work

Implications and models specificities

- Focus on particular cases: cyclic models, models with long paths
- Extension to Triadic and Triadic RCA [Bazin et al., 2024]
- Taxonomies on attributes

Connect to other tools FCA4J (`https://www.lirmm.fr/fca4j/`), RCAviz (`https://rcaviz.lirmm.fr/`), etc.

Future work

Implications and models specificities

- Focus on particular cases: cyclic models, models with long paths
- Extension to Triadic and Triadic RCA [Bazin et al., 2024]
- Taxonomies on attributes

Connect to other tools FCA4J (`https://www.lirmm.fr/fca4j/`), RCAviz (`https://rcaviz.lirmm.fr/`), etc.

Back to applications for full scale application with experts and end-users

Future work

Implications and models specificities

- Focus on particular cases: cyclic models, models with long paths
- Extension to Triadic and Triadic RCA [Bazin et al., 2024]
- Taxonomies on attributes

Connect to other tools FCA4J (`https://www.lirmm.fr/fca4j/`), RCAviz (`https://rcaviz.lirmm.fr/`), etc.

Back to applications for full scale application with experts and end-users

- Knomana

Future work

Implications and models specificities

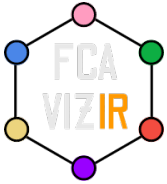
- Focus on particular cases: cyclic models, models with long paths
- Extension to Triadic and Triadic RCA [Bazin et al., 2024]
- Taxonomies on attributes

Connect to other tools FCA4J (`https://www.lirmm.fr/fca4j/`), RCAviz (`https://rcaviz.lirmm.fr/`), etc.

Back to applications for full scale application with experts and end-users

- Knomana
- Software Product Lines
- Digital debate

¡Gracias por su atención!



https://fcavizir.lirmm.fr/

Implication files prepared with https://www.lirmm.fr/fca4j/

#DigitAg
ANR Program Investments for the Future
Grant ANR-16-CONV-0004 anr

ANR SmartFCA project
Grant ANR-21-CE23-0023 anr

Baldonado, M. Q. W., Woodruff, A., and Kuchinsky, A. (2000).
Guidelines for using multiple views in information visualization.
In Proceedings of the working conference on Advanced visual interfaces (AVI), pages 110–119. ACM Press.

Bazin, A., Galasso, J., and Kahn, G. (2024).
Polyadic relational concept analysis.
Int. J. Approx. Reason., 164:109067.

Ben Sassi, I., Guenoune, H., Bazin, A., Huchard, M., Lafourcade, M., and Sallantin, J. (2024).
Dispositif d'apprentissage automatique collaboratif pour la pratique du débat.
In
JADT 2024 - 17th International Conference on Statistical Analysis of Textual Data, Bruxelles, Belgium.

Carbonnel, J., Huchard, M., and Nebut, C. (2019).

Exploring the variability of interconnected product families with relational concept analysis.
In Cetina, C., Díaz, O., Duchien, L., Huchard, M., Rabiser, R., Salinesi, C., Seidl, C., Tërnava, X., Teixeira, L., Thüm, T., and Ziadi, T., editors, Proceedings of the 23rd International Systems and Software Product Line Conference, SPLC 2019, Volume B, Paris, France, September 9-13, 2019, pages 90:1–90:8. ACM.

Duquenne, V. (2013).
Contextual implications between attributes and some representation properties for finite lattices.
In Cellier, P., Distel, F., and Ganter, B., editors, Formal Concept Analysis, pages 1–27, Berlin, Heidelberg. Springer Berlin Heidelberg.

Guigues, J. L. and Duquenne, V. (1986).
Familles minimales d'implications informatives résultant d'un tableau de données binaires.

Mathématiques et sciences humaines, 95:5–18.


Mahrach, L., Gutierrez, A., Huchard, M., Keip, P., Marnotte, P., Silvie, P., and Martin, P. (2021).
Combining implications and conceptual analysis to learn from a pesticidal plant knowledge base.
In 26th Int. Conf. on Conceptual Structures (ICCS), volume 12879 of LNCS, pages 57–72. Springer.


Roberts, J. C. (2007).
State of the art: Coordinated & multiple views in exploratory visualization.
In Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV), pages 61–71.


Shneiderman, B. (1996).
The eyes have it: A task by data type taxonomy for information visualizations.
In Proceedings of the 1996 IEEE Symposium on Visual Languages, pages 336–343. IEEE Computer Society.


Silvie, P. J., Martin, P., Huchard, M., Keip, P., Gutierrez, A., and Sarter, S. (2021).
Prototyping a knowledge-based system to identify botanical extracts for plant health in sub-saharan africa.
Plants, 10(5).