



HAL
open science

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

► **To cite this version:**

Lola Musslin, Alexandre Bazin, Marianne Huchard, Pierre Martin, Pascal Poncelet, et al.. FCAvizIR: Exploring Relational Data Set's Implications Using Metrics and Topics. CONCEPTS 2024 - 1st International Joint Conference on Conceptual Knowledge Structures, Universidad de Cádiz, Sep 2024, Cadiz, Spain. pp.132-148, 10.1007/978-3-031-67868-4_10 . lirmm-04668541

HAL Id: lirmm-04668541

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-04668541v1>

Submitted on 6 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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



@

Univ. Montpellier



, CNRS



, CIRAD



, Univ. Paul-Valéry Montpellier 3



CONCEPTS'24

Table of Contents

- 1 Introduction
- 2 RCA implications
- 3 Requirements
- 4 FCAvizIR
- 5 Conclusion

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

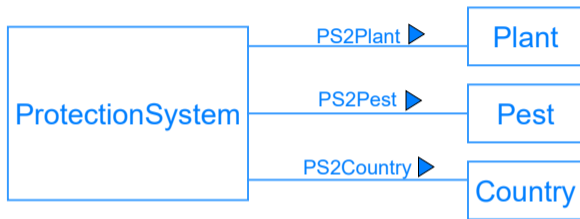
- Environmental domain (Knomana): Understanding relations between therapeutic plants, pests (bioagressors), and affected organisms to find or hypothesize solutions in phytotherapy [Mahrach et al., 2021]
- Software Engineering domain: Extracting Variability in Software Product Lines, in a relational context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

Table of Contents

- 1 Introduction
- 2 RCA implications**
- 3 Requirements
- 4 FCAvizIR
- 5 Conclusion

RCA in a nutshell

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

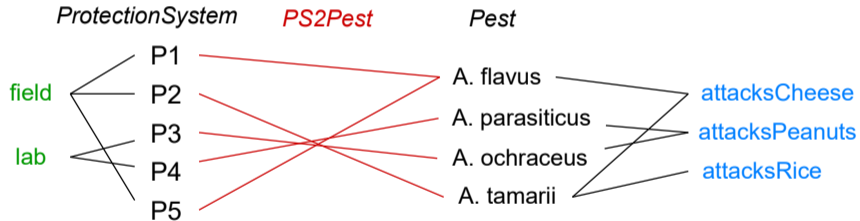


Excerpt from Knomana [Silvie et al., 2021]

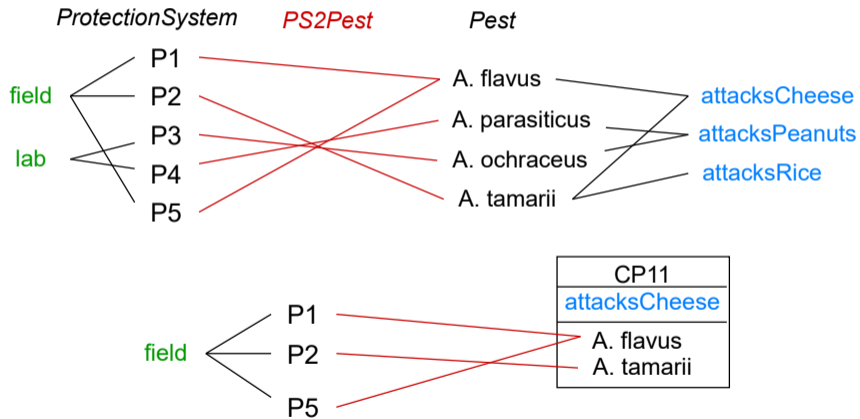
*P1 (Protection System in the field) developed in Namibia (Country):
Cinnamomum zeylanicum (Plant) is used to control Aspergillus flavus (Pest)*

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

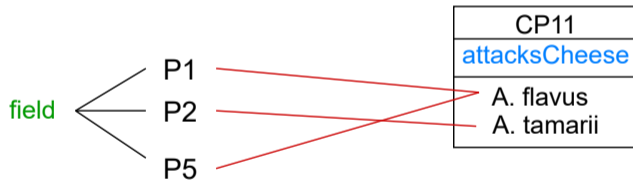
From the dataset to Pest concept CP11



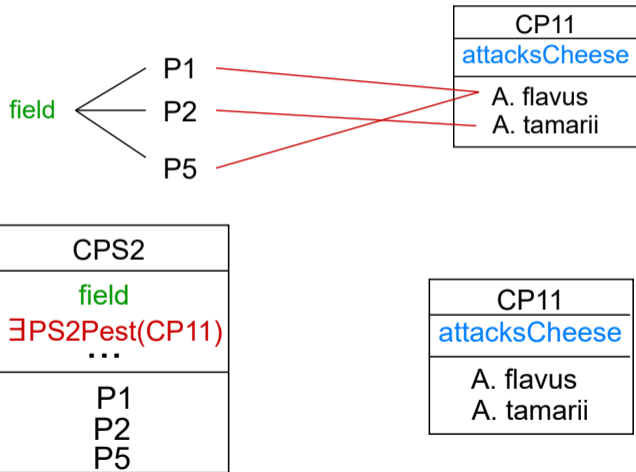
From the dataset to Pest concept CP11



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



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



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

CPS2
field \exists PS2Pest(CP11) ...
P1 P2 P5

CP11
attacksCheese
A. flavus A. tamarii

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

CPS2
field \exists PS2Pest(CP11) ...
P1 P2 P5

CP11
attacksCheese
A. flavus A. tamarii

CPS2
field \exists PS2Pest(attacksCheese) ...
P1 P2 P5

CP11
attacksCheese
A. flavus A. tamarii

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

ProtSyst	field	lab	\exists PS2Pest(attacksRice)	\exists PS2Pest(attacksCheese)	\exists PS2Pest(attacksPeanuts)	\exists PS2Pest(aspergillus)	\exists PS2Country(australAfrica)	\exists PS2Country(westernAfrica)	\exists PS2Plant(contraceptive)	\exists PS2Plant(antidysenteric)	\exists PS2Plant(lauraceae*)	\exists PS2Plant(asteraceae**)	\exists 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]:

$\exists PS2Pest(aspergillus), \exists PS2Plant(lauraceae \& comestible \& applicOil), \exists PS2Plant(aromatic \& evergreen)$
 $\Rightarrow field, \exists PS2Pest(attacksCheese), \exists PS2Country(australAfrica)$

$\exists PS2Pest(aspergillus), \exists PS2Country(westernAfrica), \exists PS2Plant(aromatic \& evergreen) \Rightarrow$
 $lab, \exists PS2Pest(attacksPeanuts), \exists PS2Plant(asteraceae \& toxic \& applicEssentialOil \& applicExtract)$

$\exists PS2Pest(attacksPeanuts), \exists PS2Pest(aspergillus), \exists PS2Plant(aromatic \& evergreen) \Rightarrow$
 $lab, \exists PS2Country(westernAfrica), \exists PS2Plant(asteraceae \& toxic \& applicEssentialOil \& applicExtract)$

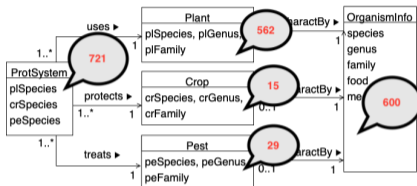
(...)

Table of Contents

- 1 Introduction
- 2 RCA implications
- 3 Requirements**
- 4 FCAvizIR
- 5 Conclusion

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.

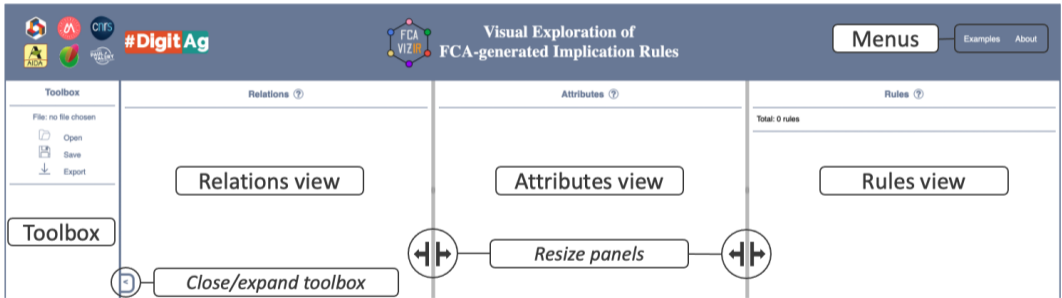
Requirements

- **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

- 1 Introduction
- 2 RCA implications
- 3 Requirements
- 4 FCAvizIR**
- 5 Conclusion

Overview

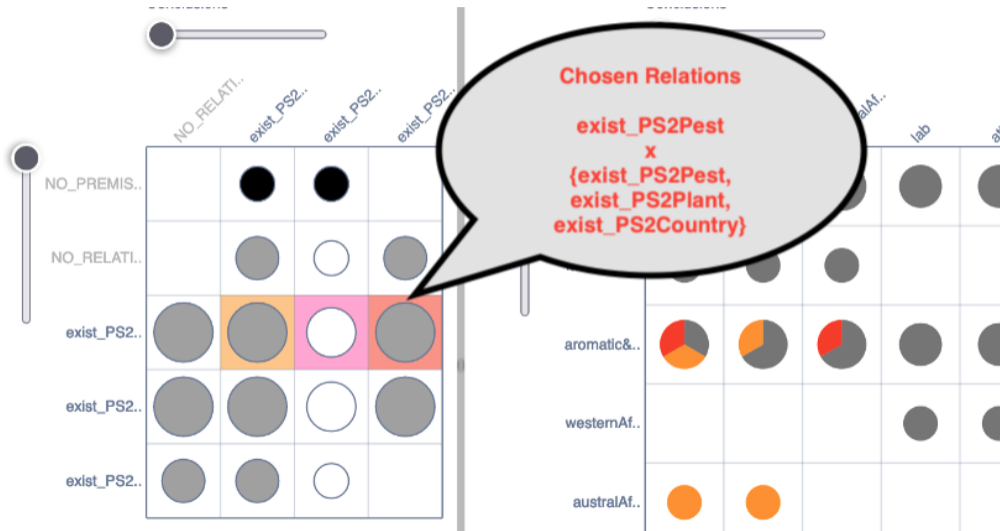


Toolbox / Parameters

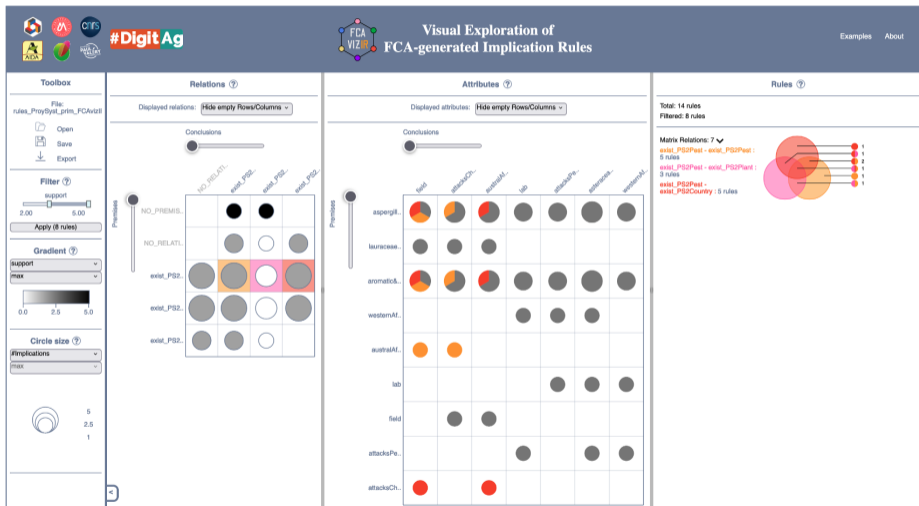
The screenshot displays the FCAvizIR software interface with several callout boxes highlighting key features:

- File:** rules_Proysyst_prim_FCAvizIR
 - Open
 - Save
 - Export
- Filter:** support (range 2.00 to 5.00) with an **Apply (8 rules)** button.
- Gradient:** support (dropdown) and max (dropdown) with a gradient bar (range 0.0 to 5.0).
- Circle size:** #Implications (dropdown) and max (dropdown) with a circle size selector (options: 5, 2.5, 1).
- Relations:** Hide empty Rows/Columns (checkbox checked), Selected Attributes, All Attributes.
- Matrix:** A matrix visualization showing relationships between attributes. The matrix is a 5x4 grid with rows labeled NO_PREMIS., NO_RELATI., exist_PS2., exist_PS2., and exist_PS2.. The columns are labeled NO_RELATI., exist_PS2., exist_PS2., and exist_PS2.. A callout box highlights the cell at row 3, column 2, labeled **exist_PS2Pest**.

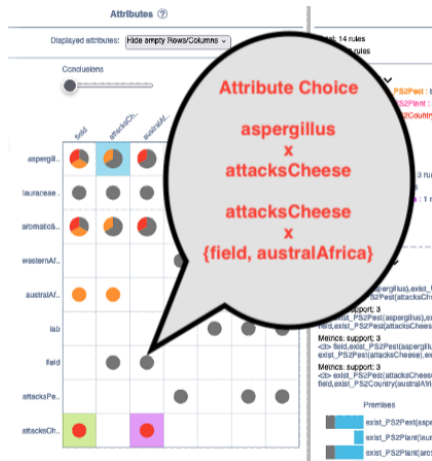
Relation Selection




Relation Selection (result)



Attribute Selection



Attribute Selection (result)



Visual Exploration of FCA-generated Implication Rules

[Examples](#) [About](#)

Toolbox

File:
rules_ProjGst_prim_FCAvizIR

Open
Save
Export


Filter

support
2.00 — 5.00

Apply (8 rules)

Gradient


support
max



0.0 2.5 5.0

Circle size

Implications
max



5

2.5

1

Relations

Displayed relations: Hide empty Rows/Columns

Conclusions

NO_REL.ATL. exist_PS2. exist_PS2.

Conclusions	NO_REL.ATL.	exist_PS2.	exist_PS2.
NO_PREMIS.	●	●	
NO_RELATL.	●	○	●
exist_PS2.	●	●	●
exist_PS2.	●	○	●
exist_PS2.	●	○	●

Attributes

Displayed attributes: Hide empty Rows/Columns

Conclusions

Conclusions	field	attacksCh.	australAF.	lab	attacksPe.	attacksCh.
aspergill.	●	●	●	●	●	●
lauraceae.	●	●	●			
aromatick.	●	●				
westernAF.			●	●	●	
australAF.	●	●				
lab				●	●	●
field		●	●			
attacksPe.				●	●	●
attacksCh.	●	●				

Rules

Total: 14 rules
Filtered: 8 rules

Matrix Relations: 7

- exist_PS2Pest - exist_PS2Pest: 5 rules
- exist_PS2Pest - exist_PS2Plant: 3 rules
- exist_PS2Pest - exist_PS2Country: 5 rules

Matrix Attributes: 4

- aspergillus - attacksCheese: 3 rules
- attacksCheese - field: 1 rules
- attacksCheese - australAfrica: 1 rules

Filtered rules: 4

Metrics: support: 2

<> exist_PS2Pest(aspergillus).exist_PS2Plant(lauraceae&acombestib&applic0).exist_PS2Plant(aromatic&evergreen) => field,exist_PS2Pest(attacksCheese),exist_PS2Country(australAfrica)

Metrics: support: 3

<> exist_PS2Pest(aspergillus).exist_PS2Country(australAfrica).exist_PS2Plant(aromatic&evergreen) => field,exist_PS2Pest(attacksCheese)

Metrics: support: 3

<> field,exist_PS2Pest(aspergillus).exist_PS2Plant(aromatic&evergreen) => exist_PS2Pest(attacksCheese),exist_PS2Country(australAfrica)

Metrics: support: 3

<> exist_PS2Pest(attacksCheese).exist_PS2Pest(aspergillus).exist_PS2Plant(aromatic&evergreen) => field,exist_PS2Country(australAfrica)

Premises

- exist_PS2Pest(asperg.
- exist_PS2Plant(laura.
- exist_PS2Plant(aroma.
- exist_PS2Country(aus.
- NO_RELATION(field)
- exist_PS2Pest(attack.

Conclusions

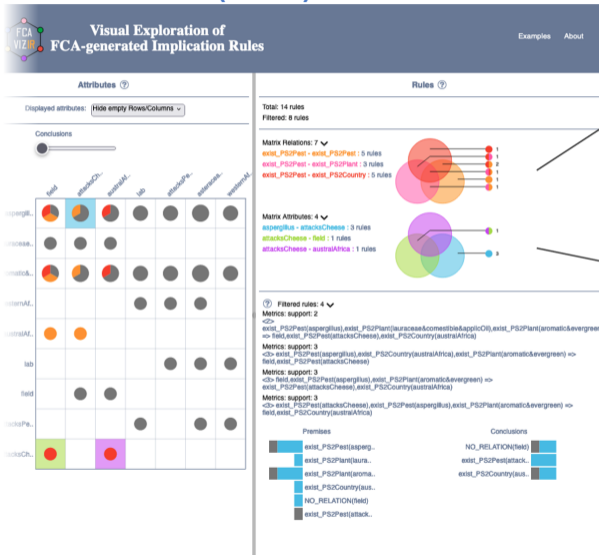
- NO_RELATION(field)
- exist_PS2Pest(attack.
- exist_PS2Country(aus.

L. Musslin et al.

CONCEPTS 2024

21 / 28

Attribute Selection (detail)

**Matrix Relations: 7**

exist_PS2Pest - exist_PS2Pest : 5 rules
 exist_PS2Pest - exist_PS2Plant : 3 rules
 exist_PS2Pest - exist_PS2Country : 5 rules

**Matrix Attributes: 4**

aspergillus - attacksCheese : 3 rules
 attacksCheese - field : 1 rules
 attacksCheese - australAfrica : 1 rules



Premise Selection (result)



Table of Contents

- 1 Introduction
- 2 RCA implications
- 3 Requirements
- 4 FCAvizIR
- 5 Conclusion**

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/`



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., Těrnava, 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\(5\)](#).