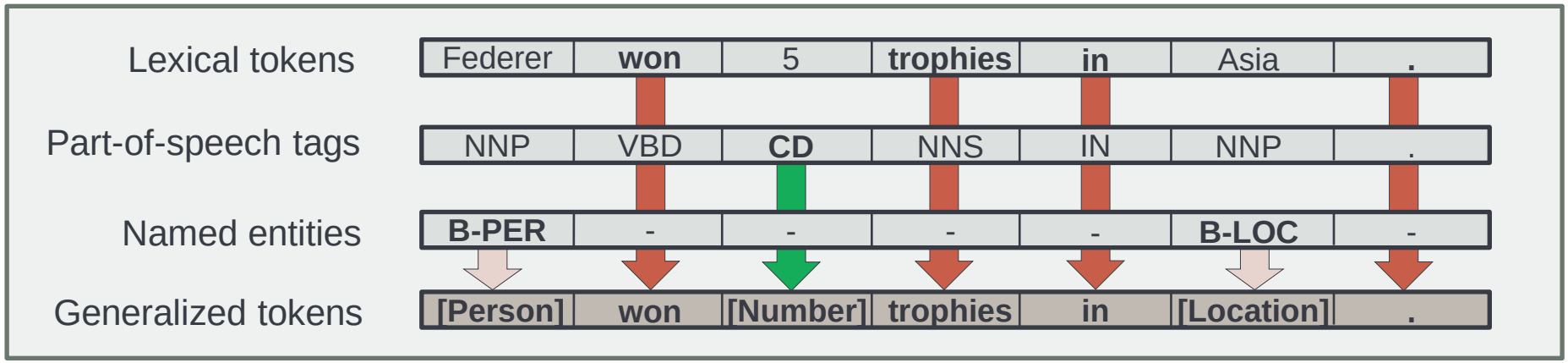
SEMI-AUTOMATIC RULE CONSTRUCTION FOR SEMANTIC LINKING OF RELATION ARGUMENTS

Janez Starc

Dunja Mladenić

Introduction

- Constructing pattern rules
  - Lexical patterns



  - Logical patterns

(numberOfTrophies [Person] [Number] [Location])

→

(numberOfTrophies RogerFederer 5 Asia)

  - Functions useful for nesting

(JuiceFn Orange-Fruit) = orange juice

Role in organization - example

Roy Blake, chief researcher and former president of Tech Lab

Role in organization - example

Roy Blake, chief researcher and former president of Tech Lab

chief [Role]

(ChiefFn [Role])

(ChiefFn Researcher)

Role in organization - example

Roy Blake, chief researcher and former president of Tech Lab

chief [Role]

former [Role]

(ChiefFn [Role])

(FormerFn [Role])

(ChiefFn Researcher)

(FormerFn President)

=

[Role]

=

[Role]

Role in organization - example

Roy Blake, chief researcher and former president of Tech Lab

Role in organization - example

Roy Blake, chief researcher and former president of Tech Lab

[Role] and [Role]

↓

(Intersection [Role] [Role])

(Intersection (ChiefFn Researcher) (FormerFn Researcher))

=

[Role]

Role in organization - example

Roy Blake, chief researcher and former president of Tech Lab

Role in organization - example

Roy Blake, chief researcher and former president of Tech Lab

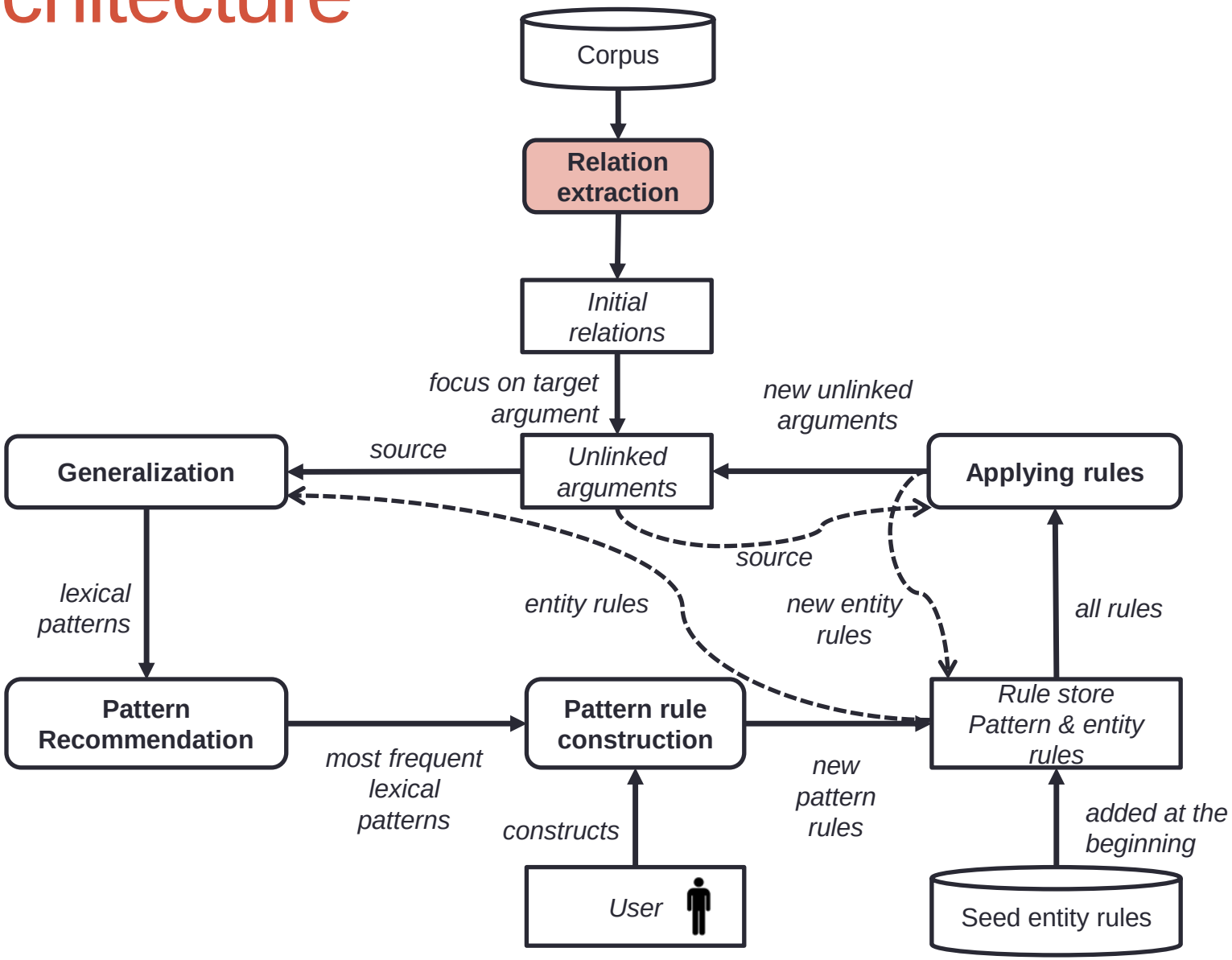
[Person], [Role] of [Organization]

↓

(roleInOrganization [Person] [Organization] [Role])

(roleInOrganization RoyBlake TechLab
(Intersection (ChiefFn Researcher) (FormerFn President)))

Architecture



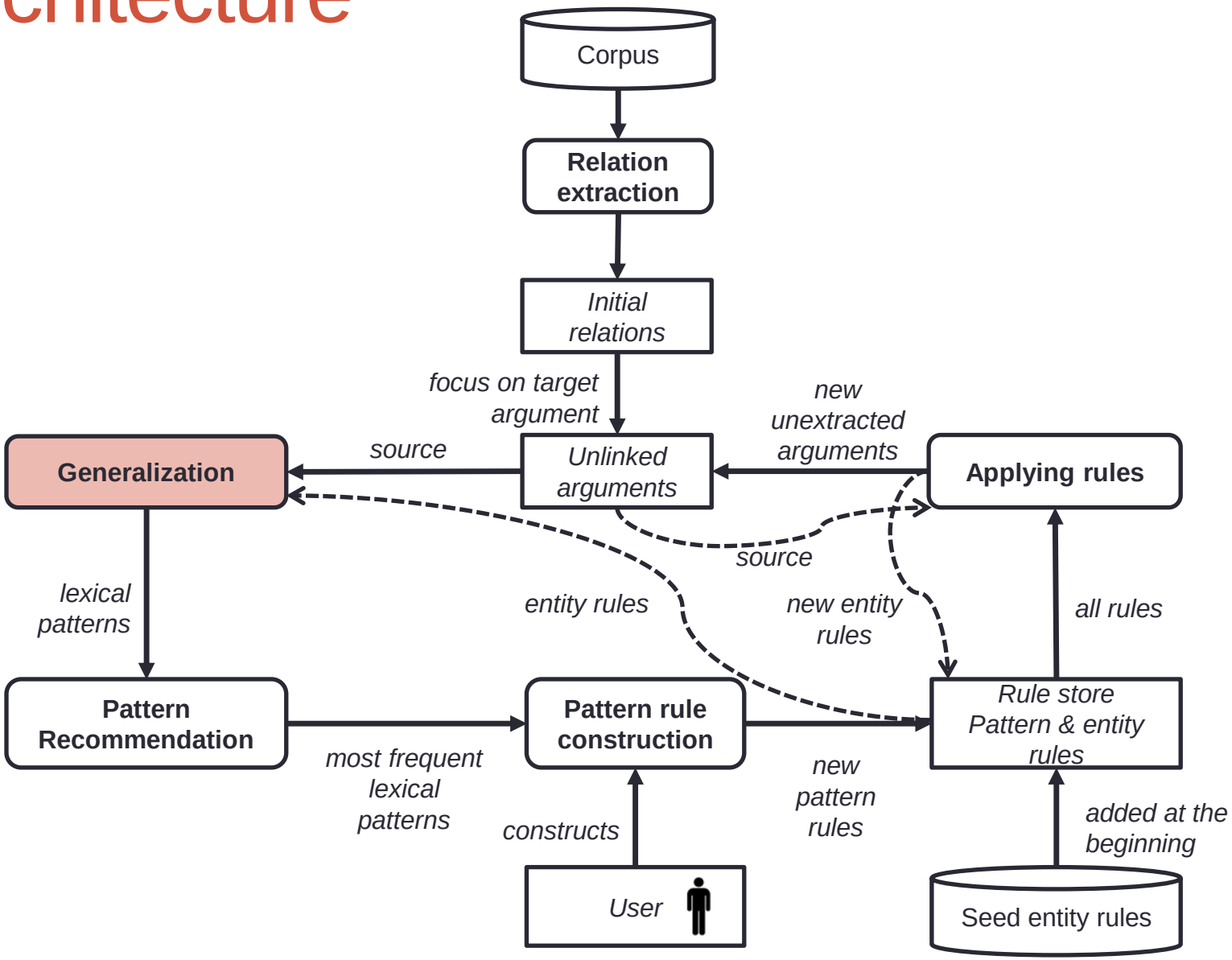
Initial relation extraction

- Applying pattern:

  [Person], [Role] of [Organization]

  - former chairman
  - manager
  - president and publisher
  - who is now president
  - who has composed the scores for most

Architecture



Generalization

chief researcher and former president

Generalization:

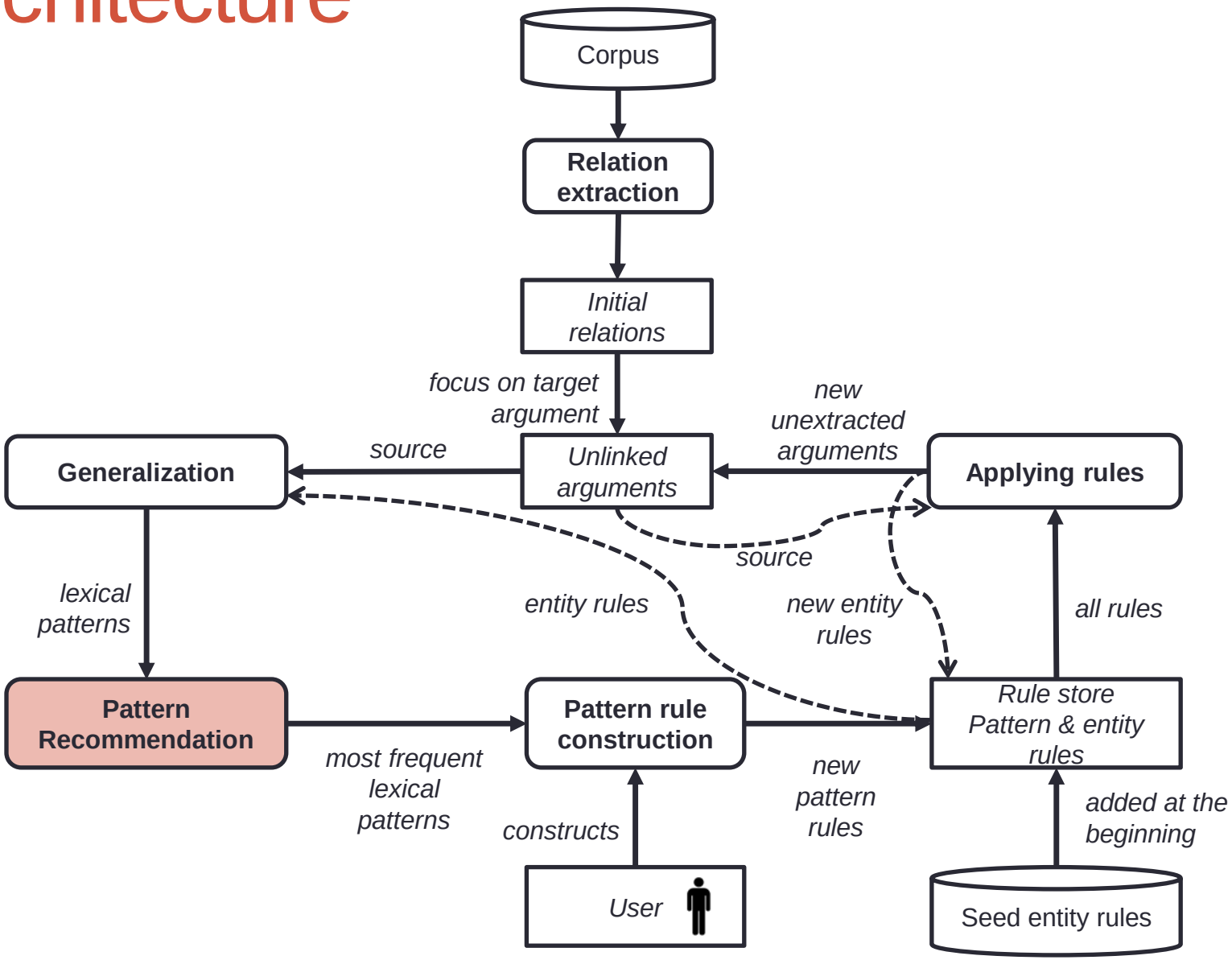
Named-entity generalization

Entity rules (seed entity rules from Freebase)

"engineer"

chief [Role] and former [Role]

Architecture

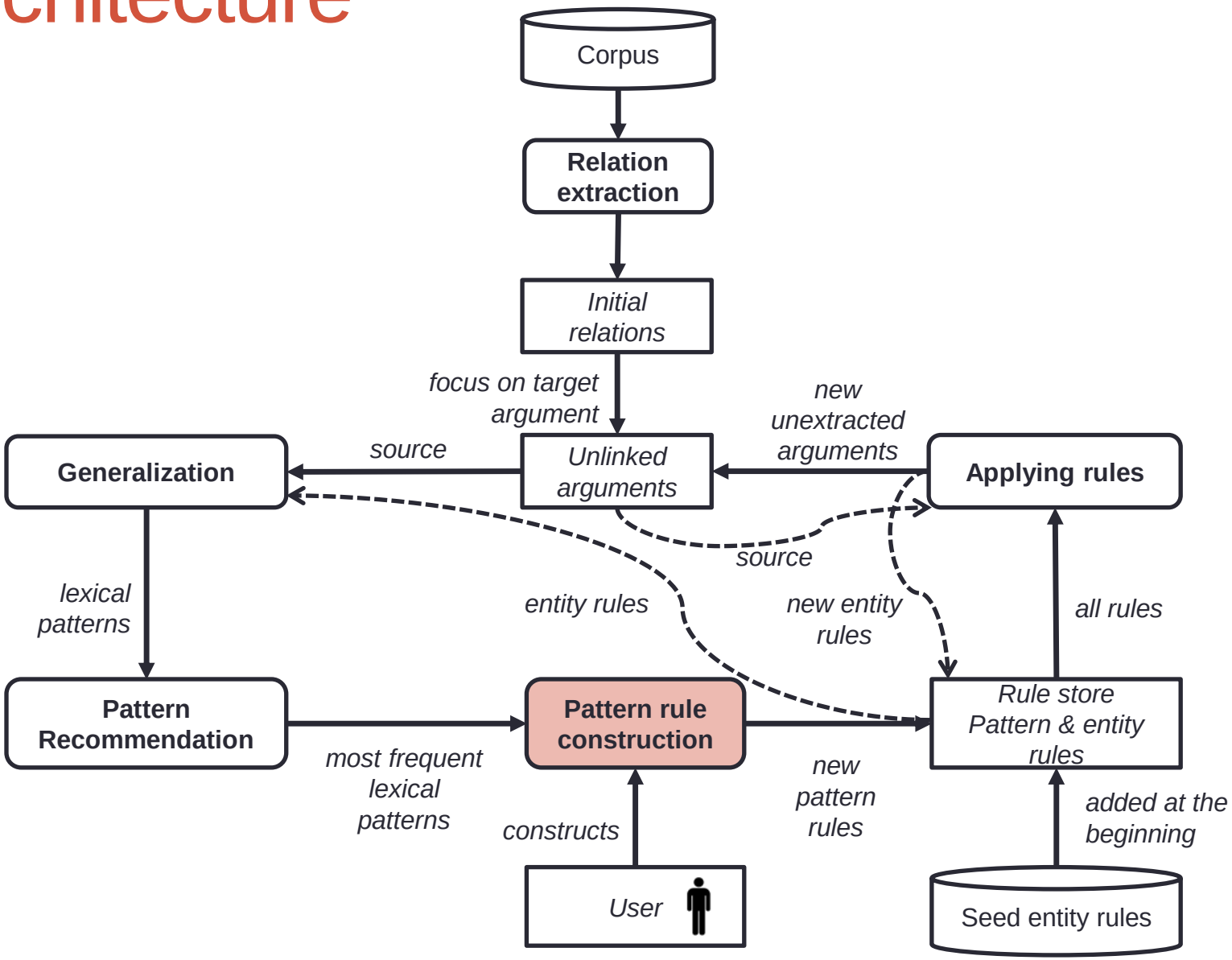


Pattern recommendation

LEXICAL PATTERN	FREQUENCY
[Role]	7149
the [Role]	1439
a [Role]	911
[Role] and [Role]	870
who is [Role]	344
former [Role]	245
…	…
[Role] and the [Role]	30

← ENTITY RULES

Architecture



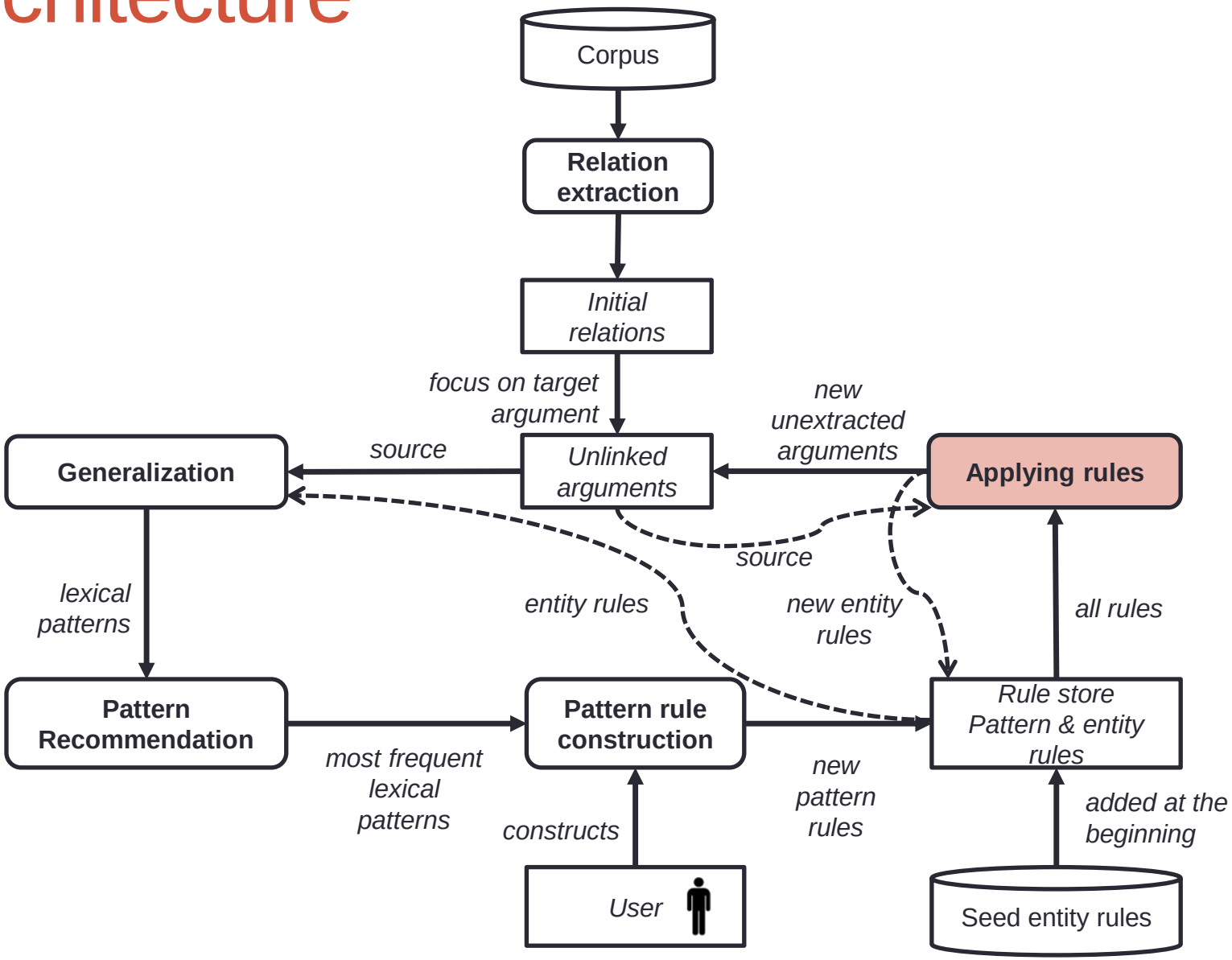
Pattern rule construction

LEXICAL PATTERN	FREQUENCY
[Role]	7149
the [Role]	1439
a [Role]	911
[Role] and [Role]	870
who is [Role]	344
former [Role]	245
…	…
[Role] and the [Role]	30

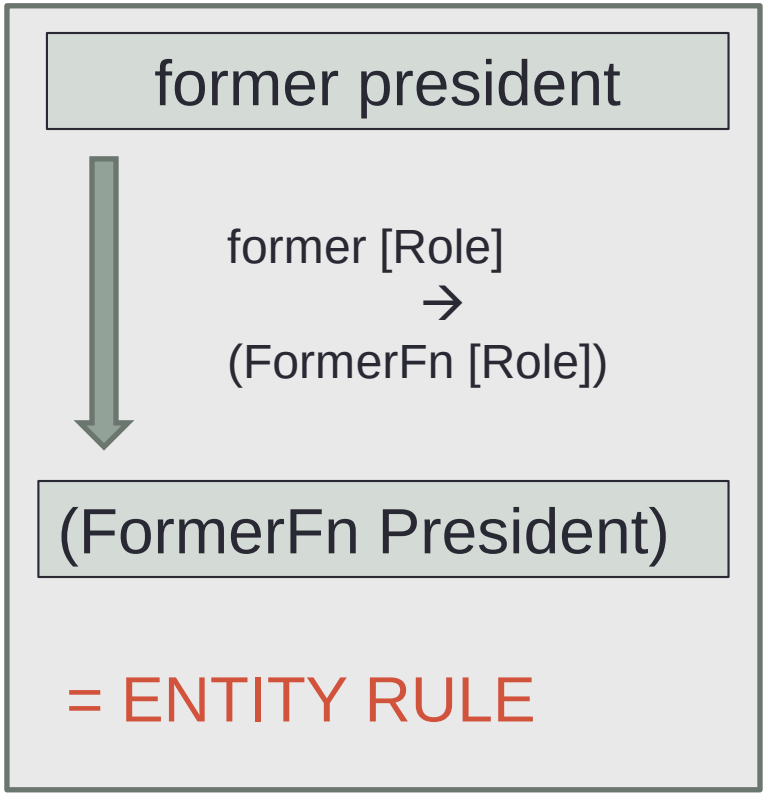
USER SELECTS

(FormerFn [Role])

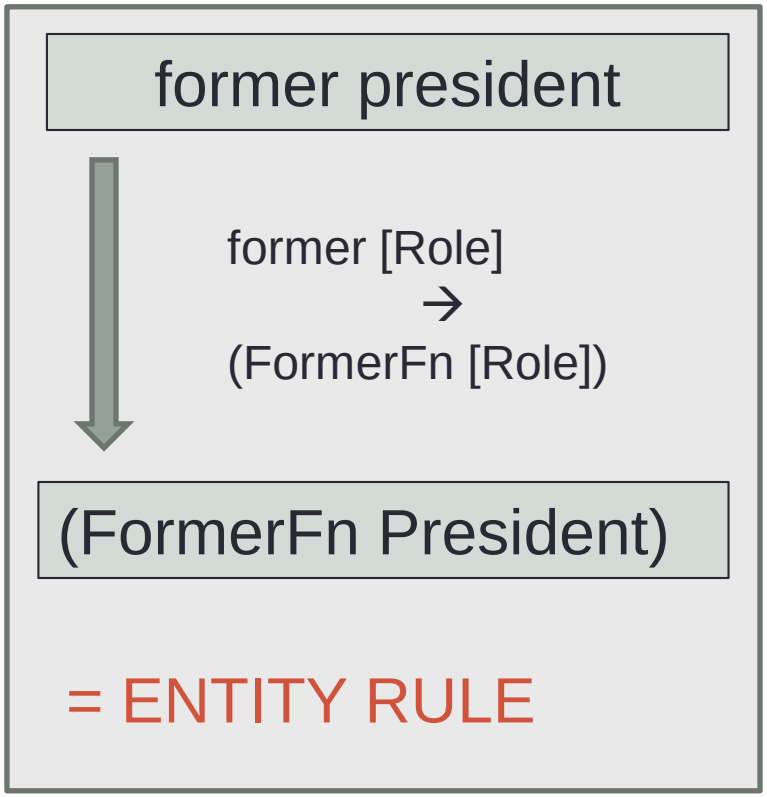
Architecture

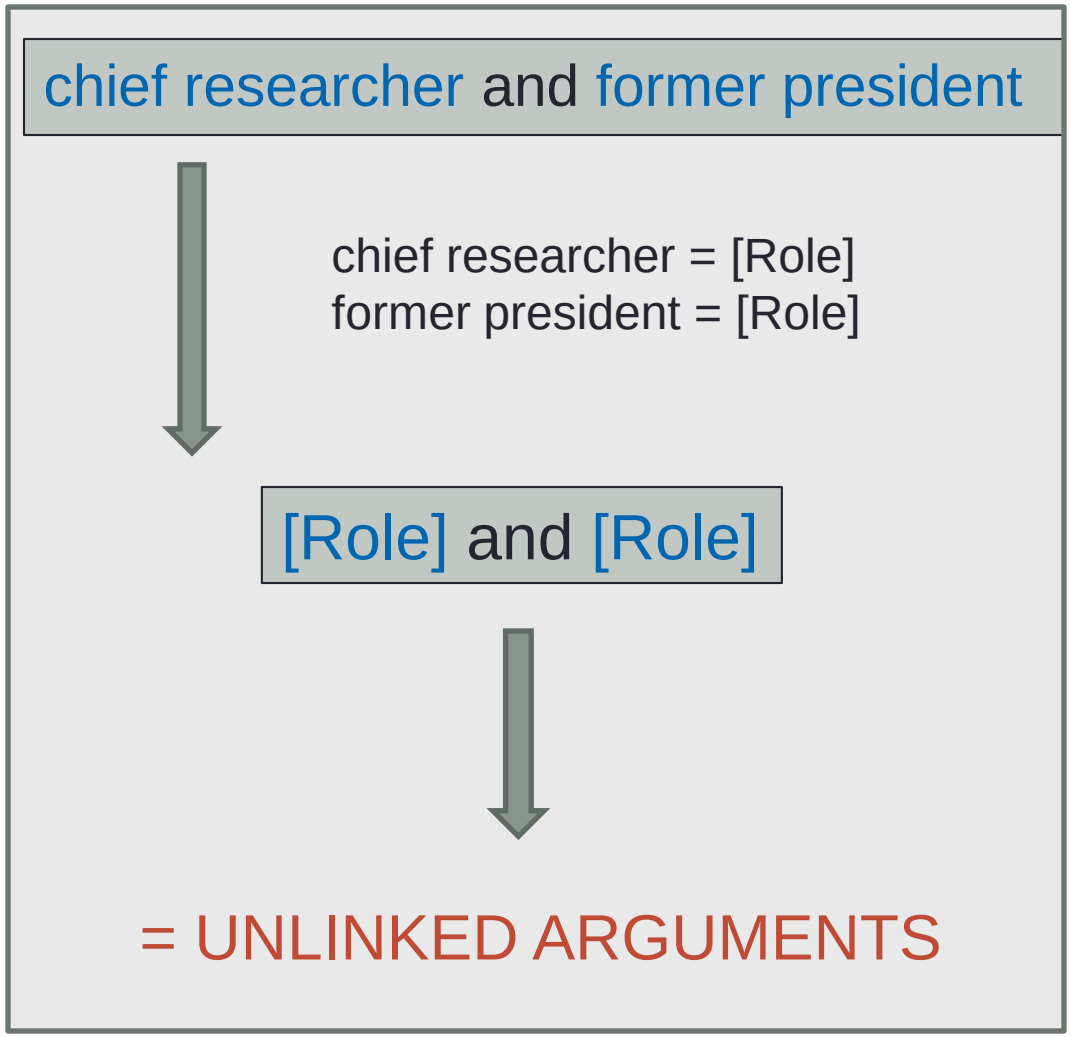


Applying pattern rules

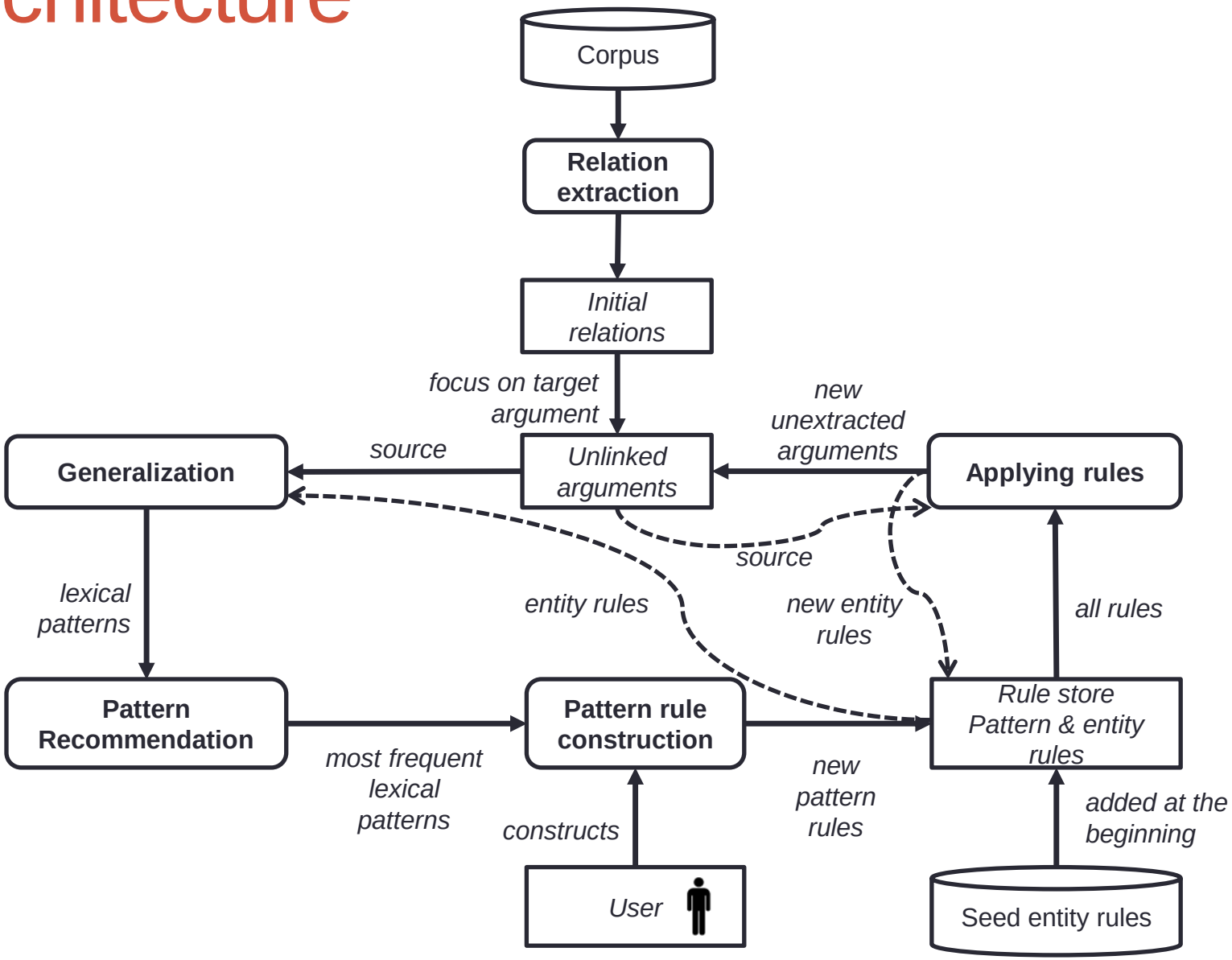


Applying pattern rules





Architecture



Next iteration

Selected on previous iteration

LEXICAL PATTERN	OLD FREQUENCY	NEW FREQUENCY
[Role]	7149	7732
~~the [Role]~~	~~1439~~	
[Role] and [Role]	870	968
a [Role]	911	911
who is [Role]	344	359
former [Role]	245	256
…	…	
~~[Role] and the [Role]~~	~~30~~	

Next iteration

Selected on previous iteration

Merge automatically

LEXICAL PATTERN	OLD FREQUENCY	NEW FREQUENCY
[Role]	7149	7732
~~the [Role]~~	~~1439~~	
[Role] and [Role]	870	968
a [Role]	911	911
who is [Role]	344	359
former [Role]	245	256
…	…	
~~[Role] and the [Role]~~	~~30~~	

Evaluation

- 1.300.000 general news articles
- 31 pattern rules created
- 14.247 completely linked roles
- 300 initial arguments left out for testing
- 100% precision, 84% recall

Conclusion and future work

- Suggesting frequent lexical patterns for linking relation arguments
- Constructing small number of rules that cover a large portion of text

- Future work
  - Automatic decomposition of text (lexical patterns)
  - Adding complexity → Wikipedia (first paragraph of person articles)

QUESTIONS ?