



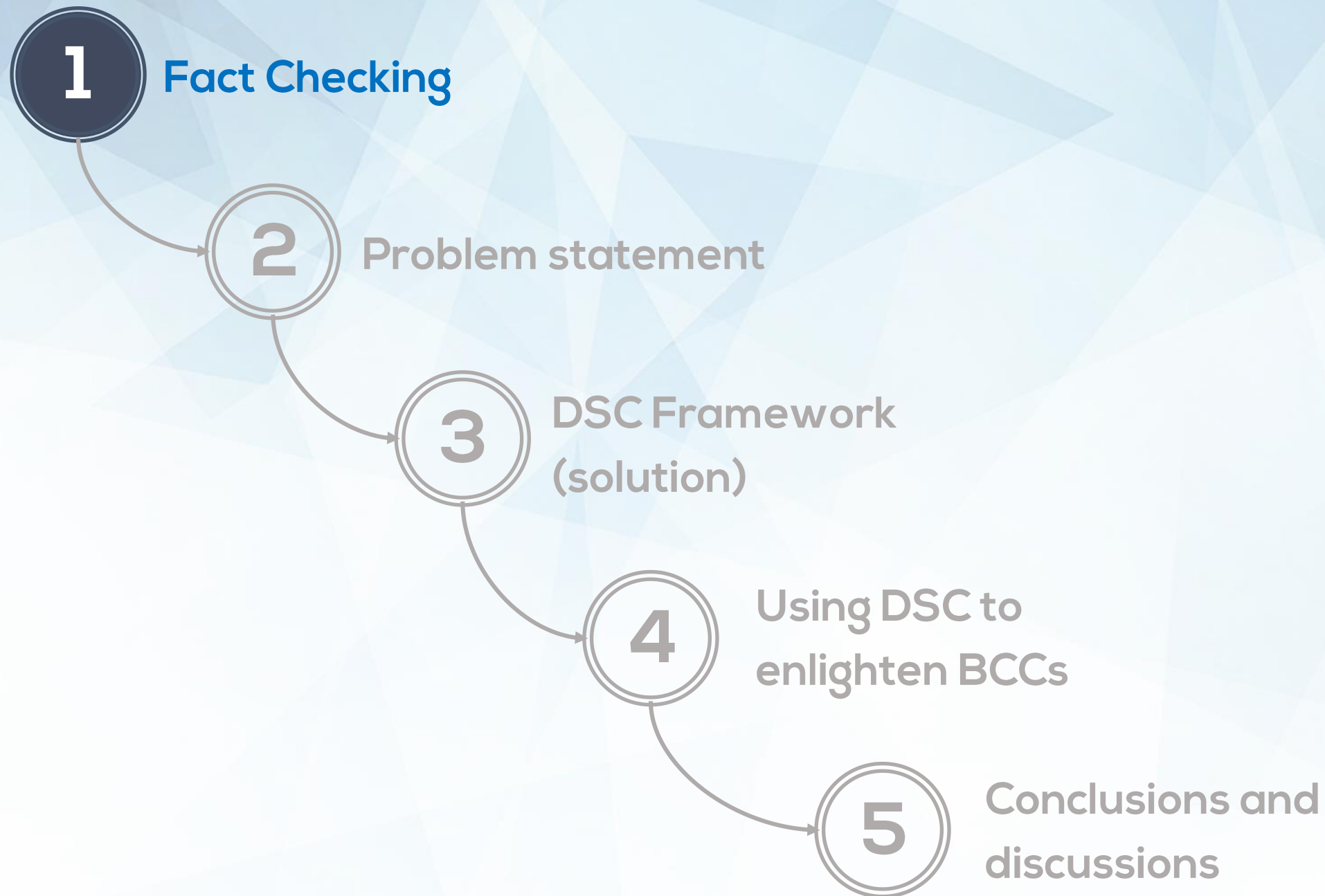
# **FLASHPOINTS:** MINING EXCEPTIONAL PAIRWISE BEHAVIOR **IN VOTE DATASETS**

*\*An exceptional model mining technique to factcheck behavioral claims*

## **AUTHORS.**

**Adnene BELFODIL** Philippe LAMARRE  
Marc PLANTEVIT Sylvie CAZALENS







**Fact checking** is the act of **checking factual assertions** in non-fictional text in order to **determine the veracity and correctness of the factual statements** in the text.



**Fact checking** is the act of **checking factual assertions** in non-fictional text in order to **determine the veracity and correctness of the factual statements** in the text.



We focus on specific type of claims namely **Behaviors Comparison Claims**.



**Fact checking** is the act of **checking factual assertions** in non-fictional text in order to **determine the veracity and correctness of the factual statements** in the text.



We focus on specific type of claims namely **Behaviors Comparison Claims**.

\* **Behaviors Comparison Claims** are statements that assert a similarity or a dissimilarity of behavior between individuals, groups, countries ...



**Fact checking** is the act of **checking factual assertions** in non-fictional text in order to **determine the veracity and correctness of the factual statements** in the text.



We focus on specific type of claims namely **Behaviors Comparison Claims**.

- \* **Behaviors Comparison Claims** are statements that assert a similarity or a dissimilarity of behavior between individuals, groups, countries ...
- \* **Several Claims** can be transformed into **BCCs** thus allowing them to be **contextualized**.

**Example:** In the European Parliament, The deputy X votes practically the same as the deputy Y.

To evaluate **at what extent this claim is valid**, several questions pop in mind implying different ways to enlighten such claims:

**Example:** In the European Parliament, The deputy X votes practically the same as the deputy Y.

To evaluate **at what extent this claim is valid**, several questions pop in mind implying different ways to **enlighten such claims**:

- Is this claim valid in the general case (when considering all votes) ? *(do the math)*



**Example:** In the European Parliament, The deputy X votes practically the same as the deputy Y.

To evaluate **at what extent this claim is valid**, several questions pop in mind implying different ways to **enlighten such claims**:

- Is this claim valid in the general case (when considering all votes) ? *(do the math)*
- Contextualize the claim:
  - Considering that the two deputies are from two different parties, are the two respective parties similar in their voting behavior ?
  - Is it **valid for every context** (time period, topic of ballots) ?

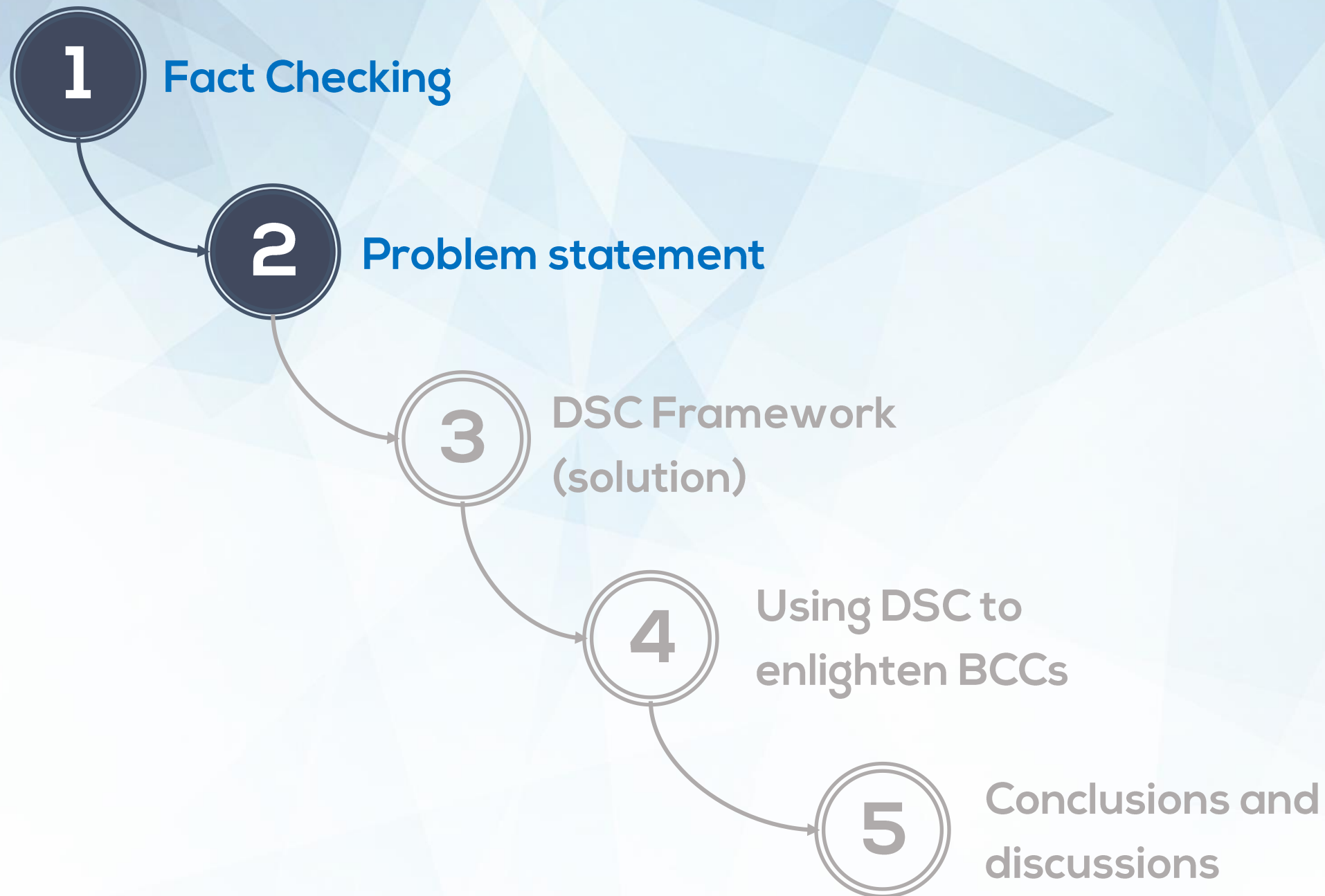
**Claim 1** In the European parliament, French deputies vote following the votes recommendation given by their respective national parties

**Claim 2** There is no national position when it comes to votes in European political group.

**Claim 3** Deputy D1 votes practically the same as a deputy D2 (Several possibilities by considering different dimensions of grouping ...)

**Claim 4** The Topic X is a hotter than the Topic Y (w.r.t. all the parliament, some countries or some political groups ...)

**Claim 5** Deputy D1 changed his behavior after 2013 compared to its national party (the two political line diverge at some point after 2013 or for particular contexts)





Voting datasets



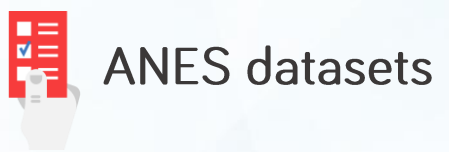
European  
parliament voting



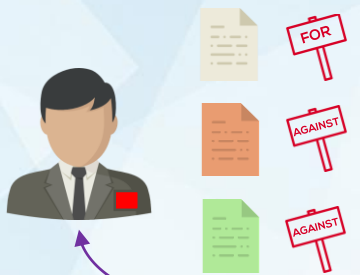
Opinion Polls



ANES datasets



## Voting datasets



Political group, country...



Theme, ballot date ...

\*context

other deputies


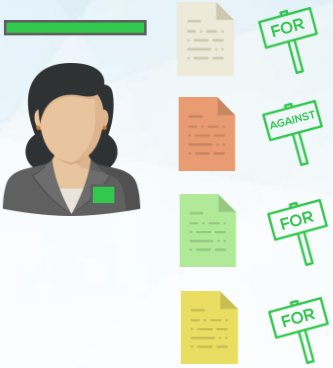


# INTRODUCTION

Left wing

Right wing

Left wing (Green)

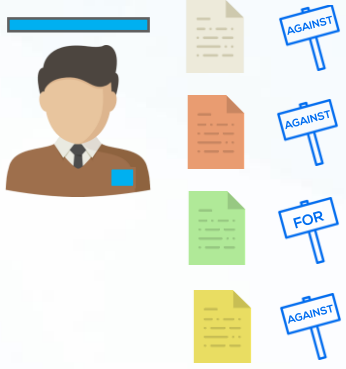
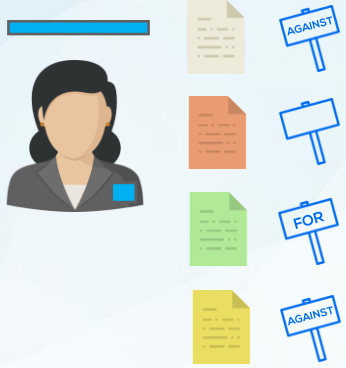


Left wing (Red)

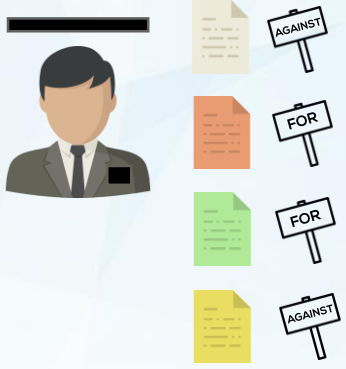
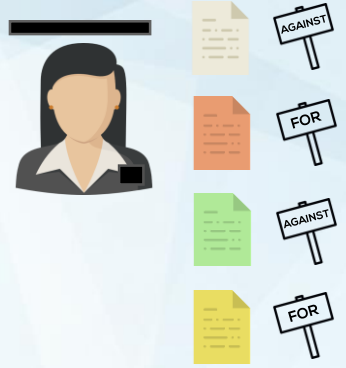


other  
deputies  
⋮  


Right wing (Blue)



Right wing (Black)



# INTRODUCTION

Left wing

Right wing


A grid showing the voting preferences of four deputies from the left wing. Each deputy is represented by a name bar, a profile icon, and a small colored square (green or red) indicating their wing affiliation. To the right of each icon are three document icons and a sign indicating their stance: 'FOR' or 'AGAINST'. The top-left deputy (female) has a green square and votes 'FOR' on all three documents. The top-right deputy (female) has a red square and votes 'AGAINST' on the first two documents and 'FOR' on the third. The bottom-left deputy (male) has a green square and votes 'AGAINST' on the first two documents and 'FOR' on the third. The bottom-right deputy (male) has a red square and votes 'FOR' on all three documents.


A grid showing the voting preferences of four deputies from the right wing. Each deputy is represented by a name bar, a profile icon, and a small colored square (blue or black) indicating their wing affiliation. To the right of each icon are three document icons and a sign indicating their stance: 'FOR' or 'AGAINST'. The top-left deputy (female) has a blue square and votes 'AGAINST' on the first two documents and 'FOR' on the third. The top-right deputy (female) has a black square and votes 'FOR' on the first two documents and 'AGAINST' on the third. The bottom-left deputy (male) has a blue square and votes 'AGAINST' on the first two documents and 'FOR' on the third. The bottom-right deputy (male) has a black square and votes 'AGAINST' on the first two documents and 'FOR' on the third.

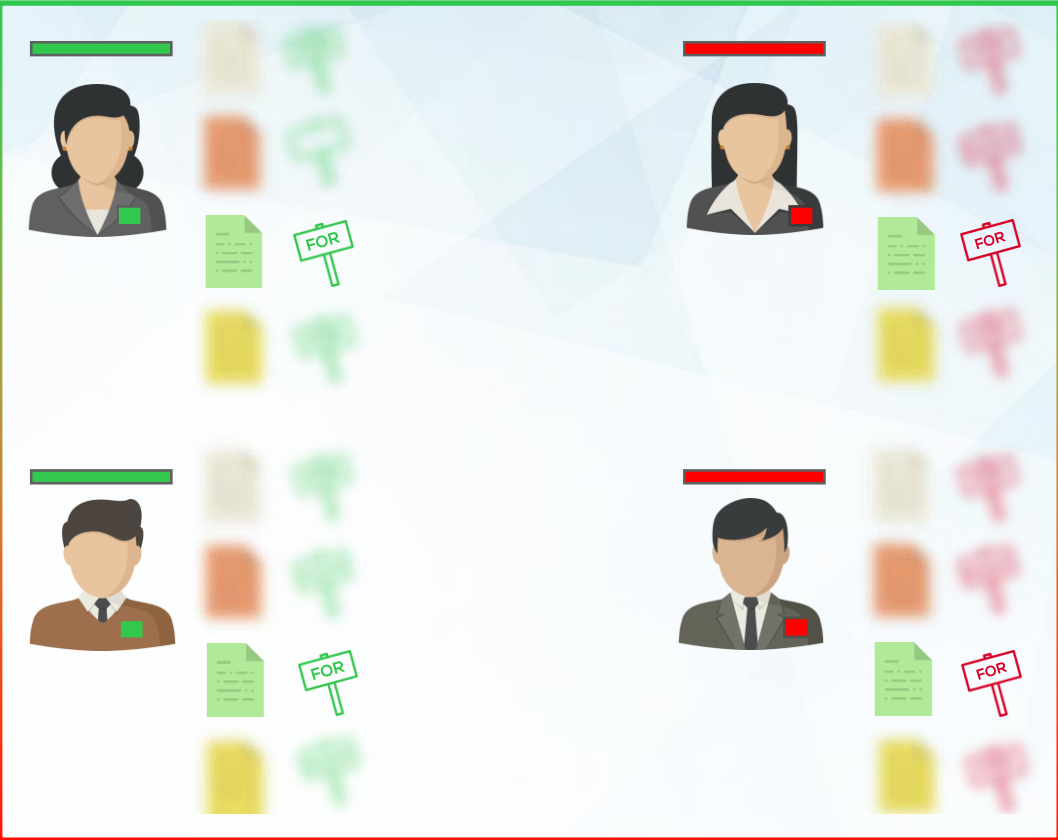
Pairwise agreement:  
25%

other  
deputies  
⋮

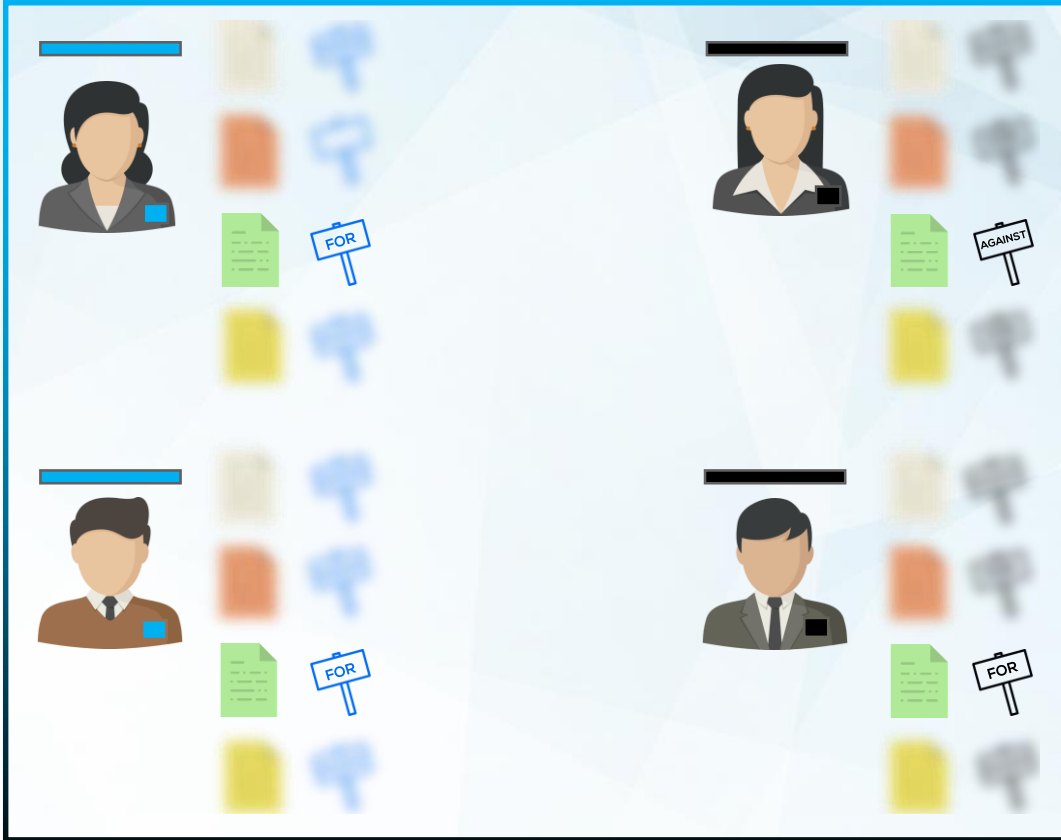
# INTRODUCTION

Left wing 

Right wing 



A 2x2 grid of voting members for the left wing. Each member is represented by a stylized icon with a colored bar above their head (green for women, red for men) and a small flag icon on their chest. To the right of each member are three document icons and a voting sign. The top-left member (woman) has a green sign that says 'FOR'. The top-right member (woman) has a red sign that says 'FOR'. The bottom-left member (man) has a green sign that says 'FOR'. The bottom-right member (man) has a red sign that says 'FOR'.



A 2x2 grid of voting members for the right wing. Each member is represented by a stylized icon with a colored bar above their head (blue for women, black for men) and a small flag icon on their chest. To the right of each member are three document icons and a voting sign. The top-left member (woman) has a blue sign that says 'FOR'. The top-right member (woman) has a black sign that says 'AGAINST'. The bottom-left member (man) has a blue sign that says 'FOR'. The bottom-right member (man) has a black sign that says 'FOR'.

 Fisheries ballots

Pairwise agreement:  
100%

other  
deputies  
⋮  




# INTRODUCTION

Left wing

Right wing

Green wing icon

FOR

AGAINST

FOR

FOR

Red wing icon

AGAINST

AGAINST

FOR

FOR

Blue wing icon

AGAINST

AGAINST

FOR

AGAINST

Black wing icon

AGAINST

FOR

AGAINST

FOR

Green wing icon

AGAINST

AGAINST

FOR

AGAINST

Red wing icon

FOR

AGAINST

FOR

FOR

Blue wing icon

AGAINST

AGAINST

FOR

AGAINST

Black wing icon

AGAINST

FOR

FOR

AGAINST

# INTRODUCTION



# INTRODUCTION





We introduce the problem of discovering **particular contexts** and **collections of individuals** such that their **pairwise behavior** exceptionally differs from their usual one



We introduce the problem of discovering **particular contexts** and **collections of individuals** such that their **pairwise behavior** exceptionally differs from their usual one



Find the **top-k three-set patterns**  $(c, g', g'')$  w.r.t. some quality measure  $\varphi$



We introduce the problem of discovering **particular contexts** and **collections of individuals** such that their **pairwise behavior** exceptionally differs from their usual one



Find the **top-k three-set patterns**  $(c, g', g'')$  w.r.t. some quality measure  $\varphi$

Context

definition by **intent** of a **subset of items**



We introduce the problem of discovering **particular contexts** and **collections of individuals** such that their **pairwise behavior** exceptionally differs from their usual one



Find the **top-k three-set patterns**  $(\underline{c}, g', g'')$  w.r.t. some quality measure  $\varphi$

Context

definition by **intent** of a **subset of items**

$g' \& g''$

definition by **intent** of a **subset of individuals**



We introduce the problem of discovering **particular contexts** and **collections of individuals** such that their **pairwise behavior** exceptionally differs from their usual one



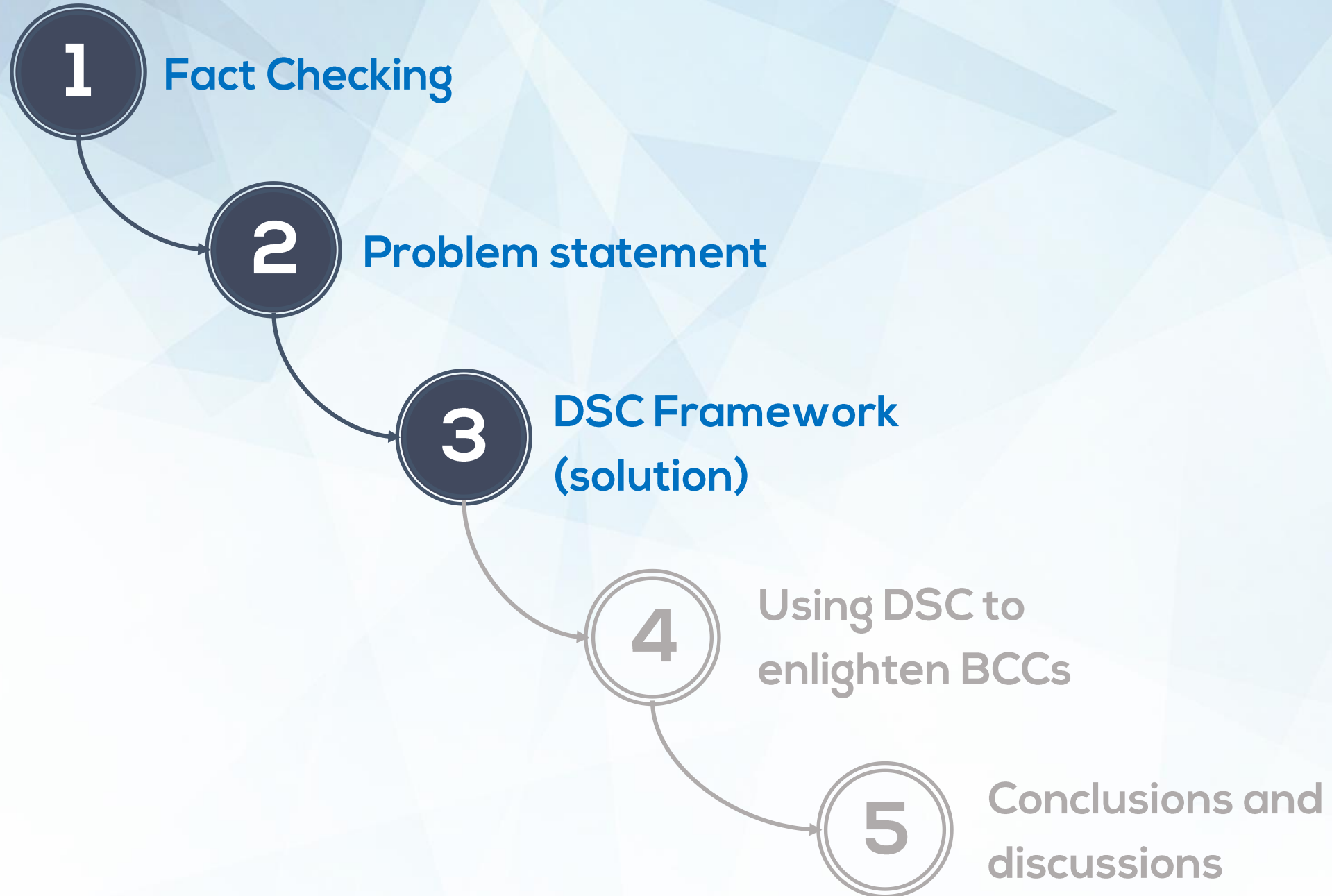
Find the **top-k three-set patterns**  $(c, g', g'')$  w.r.t. some quality measure  $\varphi$

Example:

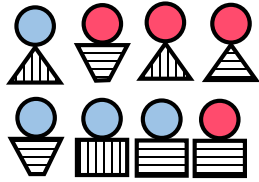
(Consumer Protection in General Ballots voted in between 2015 and 2016 , German Deputies, Italian Deputies)

We observe a **significant decrease of pairwise agreement**





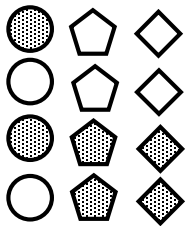
# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK



Reviewers  
*(eg. Users, Deputies)*

	○	■	△	...
●	⌞	⌞		
●	⌞			
●	⌞	⌞	⌞	
⋮				⋮

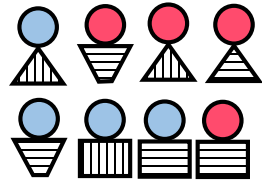
Reviews  
*(eg. Scores, Votes)*



Reviewees  
*(eg. Movies, Vote ballots)*

# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

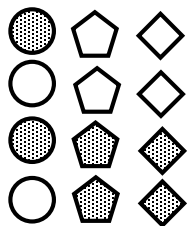
Dataset example: Parliament voting dataset



Reviewers  
(eg. Users, Deputies)

	○	◻	△	...
●	⌞	⌞		
●	⌞			
●	⌞	⌞	⌞	
⋮				

Reviews  
(eg. Scores, Votes)

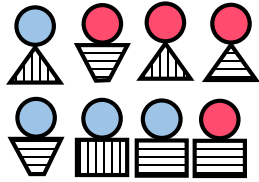


Reviewees  
(eg. Movies, Vote ballots)

<i>Items (Ballots) - E</i>			<i>Individuals (Deputies) - U</i>			<i>Outcome</i>
<i>Idsession</i>	<i>Date</i>	<i>Theme</i>	<i>Full name</i>	<i>National Party</i>	<i>Political Group</i>	<i>Vote</i>
001	2017/03/17	1.10 Justice 2.10 Europe coop	Lavrilleux	LR	PPE	For
001	2017/03/17	1.10 Justice 2.10 Europe coop	Philippot	FN	ENF	Against
002	2017/04/11	3.10 Agriculture	Lavrilleux	LR	PPE	For
002	2017/04/11	3.10 Agriculture	Philippot	FN	ENF	For
002	2017/04/11	3.10 Agriculture	Arnatu	FN	ENF	For
003	2017/04/11	1.20 Security	Le Grip	LR	PPE	Abstain

# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

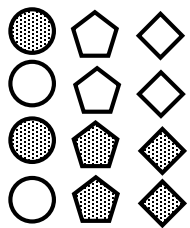
Dataset example: Parliament voting dataset



Reviewers  
(eg. Users, Deputies)

	○	■	△	...
●	⌞	⌞		
●	⌞			
●	⌞	⌞	⌞	

Reviews  
(eg. Scores, Votes)



Reviewees  
(eg. Movies, Vote ballots)

Items (Ballots) - E			Individuals (Deputies) - U			Outcome
Idsession	Date	Theme	Full name	National Party	Political Group	Vote
001	2017/03/17	1.10 Justice 2.10 Europe coop	Lavrilleux	LR	PPE	For
001	2017/03/17	1.10 Justice 2.10 Europe coop	Philippot	FN	ENF	Against
002	2017/04/11	3.10 Agriculture	Lavrilleux	LR	PPE	For
002	2017/04/11	3.10 Agriculture	Philippot	FN	ENF	For
002	2017/04/11	3.10 Agriculture	Arnatu	FN	ENF	For
003	2017/04/11	1.20 Security	Le Grip	LR	PPE	Abstain

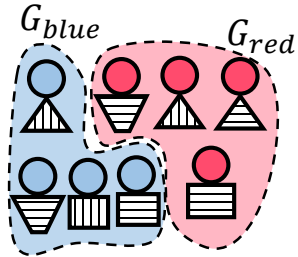
Descriptions attributes\* over  
items (context)

Descriptions attributes\* over  
individuals

\*numeric, nominal, hierarchical multi-tag attributes

# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

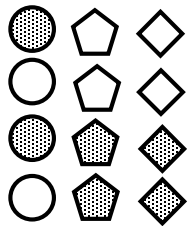
1 *Constitute groups*  
*(eg. By head color)*



Reviewers  
*(eg. Users, Deputies)*

	○	■	△	...
●	□	□		
●	□			
●	□	□	□	
⋮				⋮

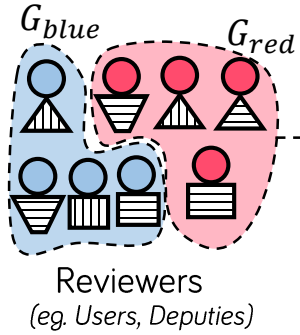
Reviews  
*(eg. Scores, Votes)*



Reviewees  
*(eg. Movies, Vote ballots)*

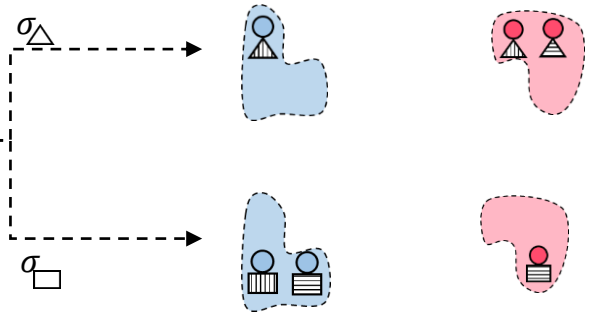
# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

**1** *Constitute groups*  
(eg. By head color)



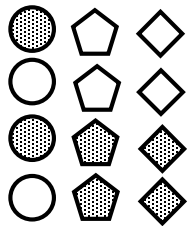
**2**

Generate a particular couple of subsets of users  
eg. Confront  $\square$  vs.  $\triangle$



	○	◻	△	...
●	◻	◻		
●	◻			
●	◻	◻	◻	
⋮				⋮

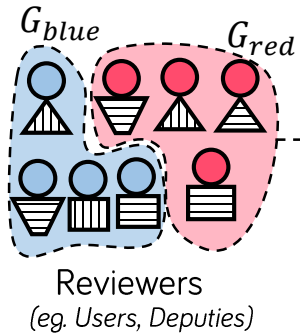
Reviews  
(eg. Scores, Votes)



Reviewees  
(eg. Movies, Vote ballots)

# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

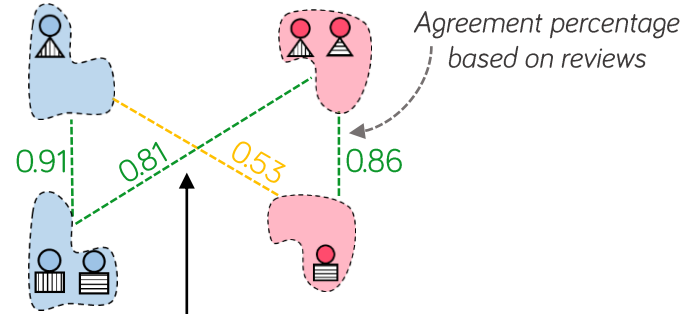
**1** *Constitute groups*  
(eg. By head color)



**2** *Generate a particular couple of subsets of users*  
eg. Confront  $\square$  vs.  $\triangle$

$\sigma_{\triangle}$   
 $\sigma_{\square}$

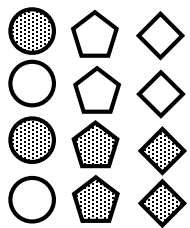
**3** *Global pairwise behavior*



Consider all reviewees

	○	◻	△	...
●	◻	△		
○	◻	△		
●	◻	△		
○	◻	△		

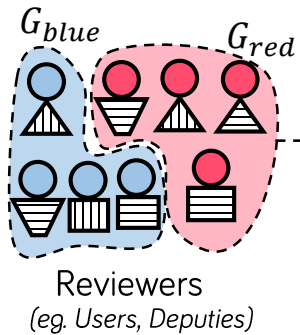
Reviews  
(eg. Scores, Votes)



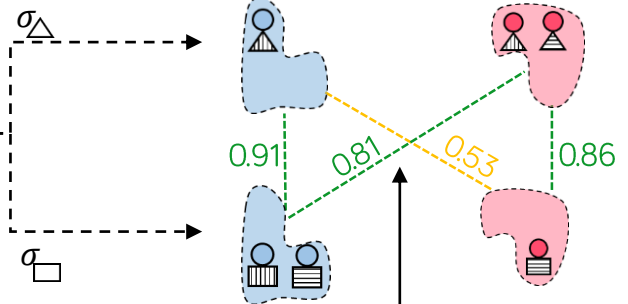
Reviewees  
(eg. Movies, Vote ballots)

# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

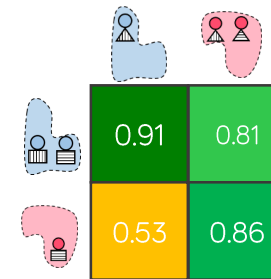
**1** *Constitute groups*  
(eg. By head color)



**2** *Generate a particular couple of subsets of users*  
eg. Confront  $\square$  vs.  $\triangle$

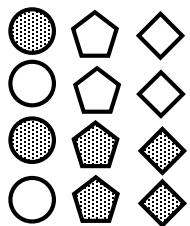


**3** *Global pairwise behavior*



	$\circ$	$\square$	$\triangle$	...
	$\uparrow$	$\uparrow$		
	$\uparrow$			
	$\uparrow$	$\uparrow$	$\uparrow$	
...				

Reviews  
(eg. Scores, Votes)



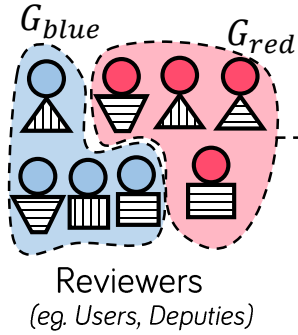
Reviewees  
(eg. Movies, Vote ballots)

Consider all reviewees



# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

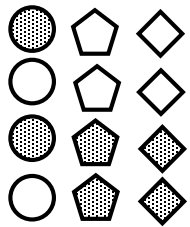
**1** Constitute groups  
(eg. By head color)



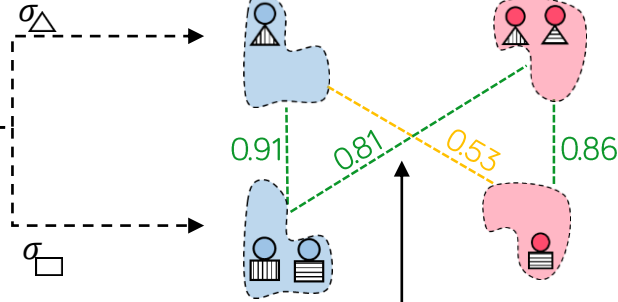
**2** Generate a particular couple of subsets of users  
eg. Confront  $\square$  vs.  $\triangle$

	○	◻	△	...
●	◻	△		
○	◻			
●	◻	△		
○	◻	△		

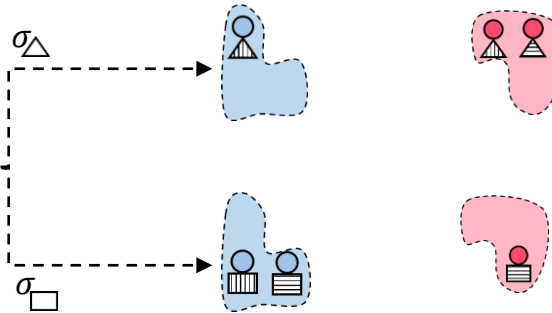
Reviews  
(eg. Scores, Votes)



**3** Global pairwise behavior



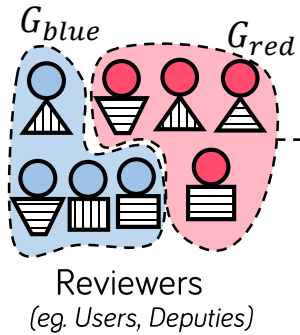
Consider all reviewees



	$G_{blue}$	$G_{red}$
$G_{blue}$	0.91	0.81
$G_{red}$	0.53	0.86

# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

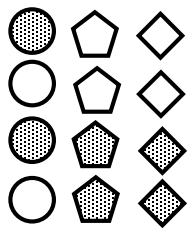
**1** *Constitute groups*  
(eg. By head color)



**2** *Generate a particular couple of subsets of users*  
eg. Confront  $\square$  vs.  $\triangle$

	○	◻	△	...
●	◻	△		
●	◻			
●	◻	△		
...				

Reviews  
(eg. Scores, Votes)

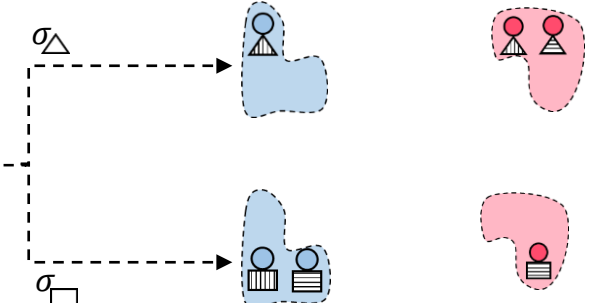
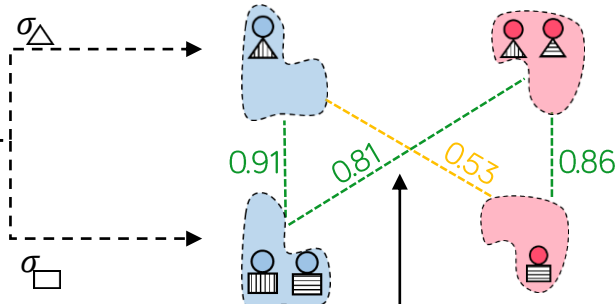


Reviewees  
(eg. Movies, Vote ballots)

Consider all reviewees

**4** *Generate a subset of reviewees*  
eg. Dotted diamonds

**3** *Global pairwise behavior*

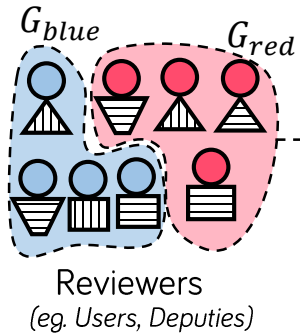


	0.91	0.81
	0.53	0.86

# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

DSC OVERVIEW

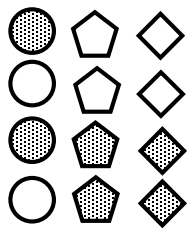
**1** *Constitute groups*  
(eg. By head color)



**2** *Generate a particular couple of subsets of users*  
eg. Confront  $\square$  vs.  $\triangle$

	$\circ$	$\square$	$\triangle$	...
$\bullet$	$\uparrow$	$\uparrow$		
$\square$	$\uparrow$			
$\triangle$	$\uparrow$	$\uparrow$		
...				

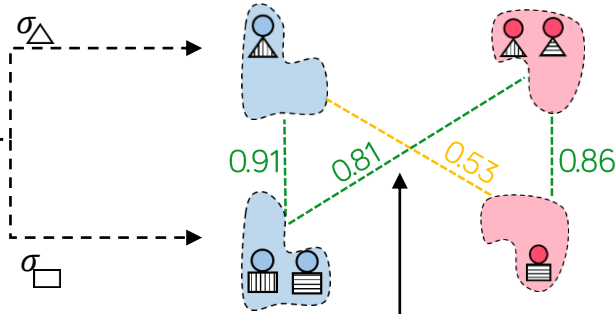
Reviews  
(eg. Scores, Votes)



**4** *Generate a subset of reviewees*  
eg. Dotted diamonds  $\blacklozenge$

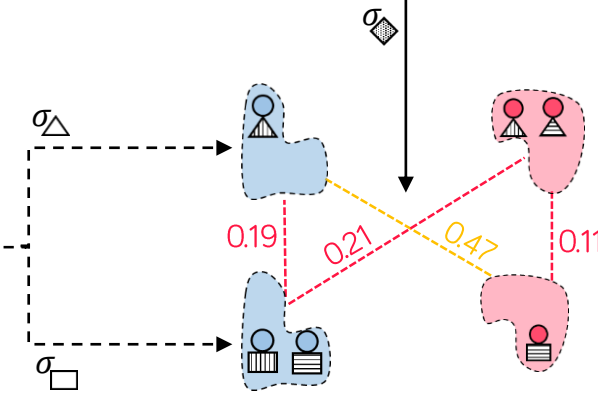
Consider all reviewees

**3** *Global pairwise behavior*



$\triangle$	$\triangle$
0.91	0.81
$\square$	$\square$
0.53	0.86

**5** *Contextual pairwise behavior*

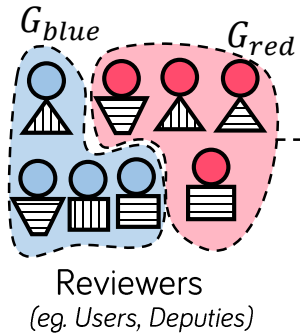


$\triangle$	$\triangle$
0.19	0.21
$\square$	$\square$
0.47	0.11

# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

DSC OVERVIEW

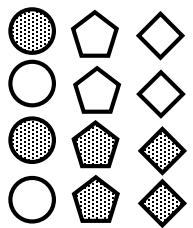
**1** *Constitute groups*  
(eg. By head color)



**2** *Generate a particular couple of subsets of users*  
eg. Confront  $\square$  vs.  $\triangle$

	$\circ$	$\square$	$\triangle$	...
$\bullet$	$\uparrow$	$\uparrow$		
$\square$	$\uparrow$			
$\triangle$	$\uparrow$	$\uparrow$		
...				

*Reviews*  
(eg. Scores, Votes)

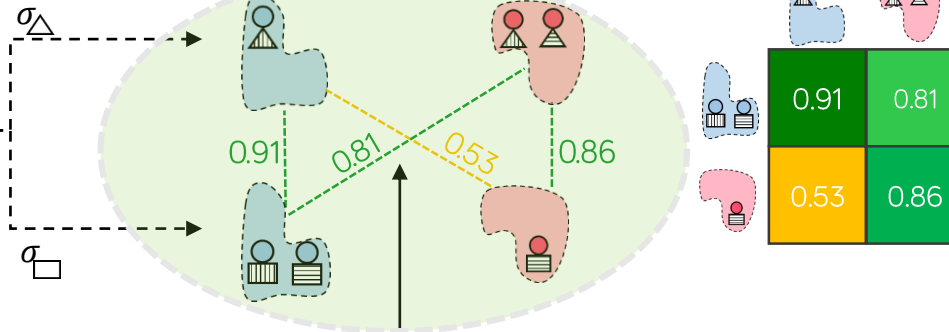


*Reviewees*  
(eg. Movies, Vote ballots)

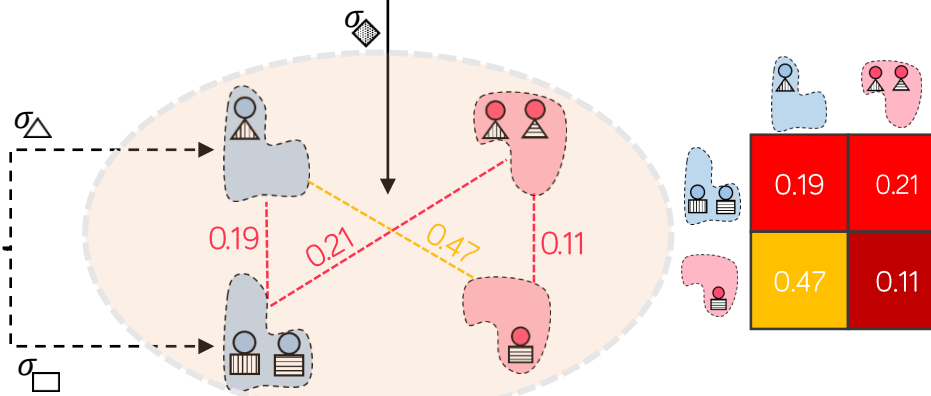
**4** *Generate a subset of reviewees*  
eg. Dotted diamonds  $\blacklozenge$

Consider all reviewees

**3** *Global pairwise behavior*

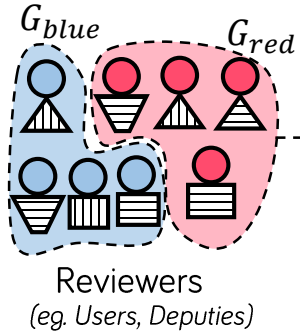


**5** *Contextual pairwise behavior*



# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

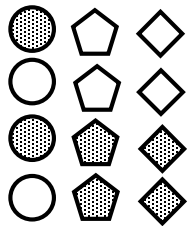
**1** *Constitute groups*  
(eg. By head color)



**2** *Generate a particular couple of subsets of users*  
eg. Confront  $\square$  vs.  $\triangle$

	$\circ$	$\square$	$\triangle$	...
$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	
$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	
$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	

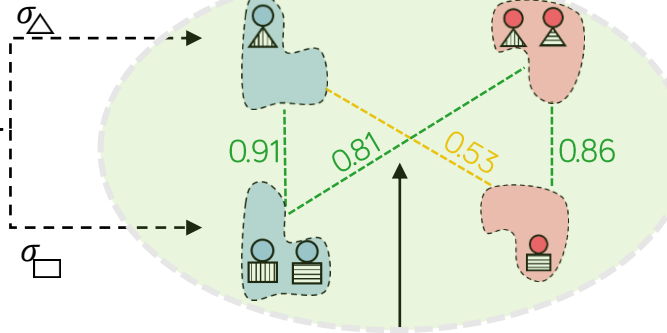
Reviews  
(eg. Scores, Votes)



**4** *Generate a subset of reviewees*  
eg. Dotted diamonds  $\blacklozenge$

Consider all reviewees

**3** *Global pairwise behavior*

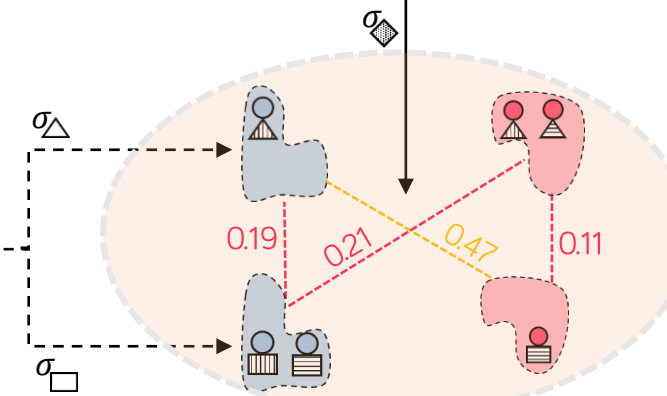


$\text{reviewer icon}$	$\text{reviewer icon}$
0.91	0.81
0.53	0.86

**6** *Compare models*  
to evaluate the intensity of changes

VS.

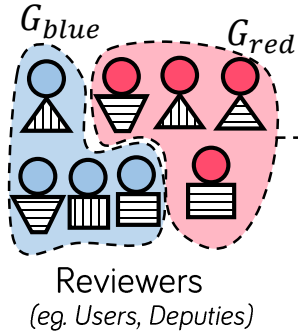
**5** *Contextual pairwise behavior*



$\text{reviewer icon}$	$\text{reviewer icon}$
0.19	0.21
0.47	0.11

# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK

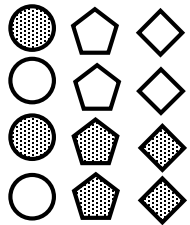
**1** Constitute groups  
(eg. By head color)



**2** Generate a particular couple of subsets of users  
eg. Confront  $\square$  vs.  $\triangle$

	$\circ$	$\square$	$\triangle$	...
$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	
$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	
$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	$\text{reviewer icon}$	

Reviews  
(eg. Scores, Votes)

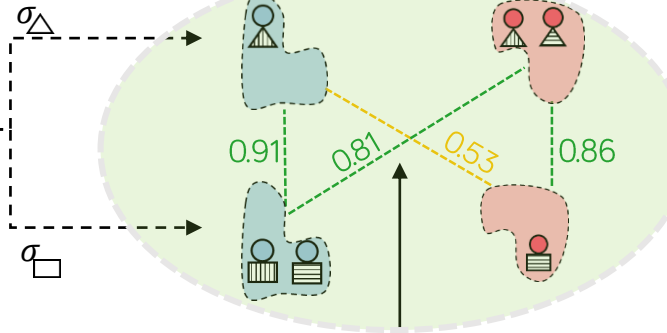


Reviewees  
(eg. Movies, Vote ballots)

**4** Generate a subset of reviewees  
eg. Dotted diamonds  $\blacklozenge$

Consider all reviewees

**3** Global pairwise behavior



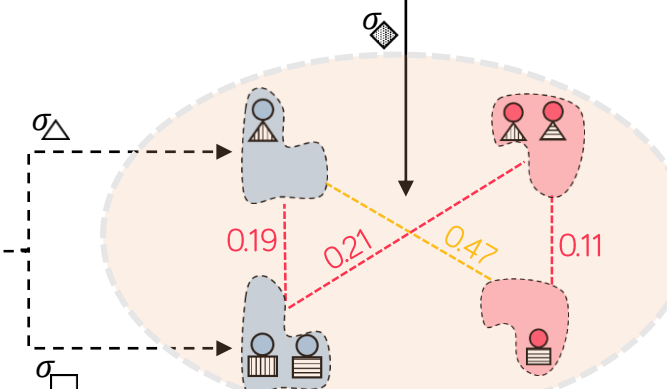
$\text{reviewer icon}$	$\text{reviewer icon}$
0.91	0.81
0.53	0.86

**6** Compare models  
to evaluate the  
intensity of changes

VS.

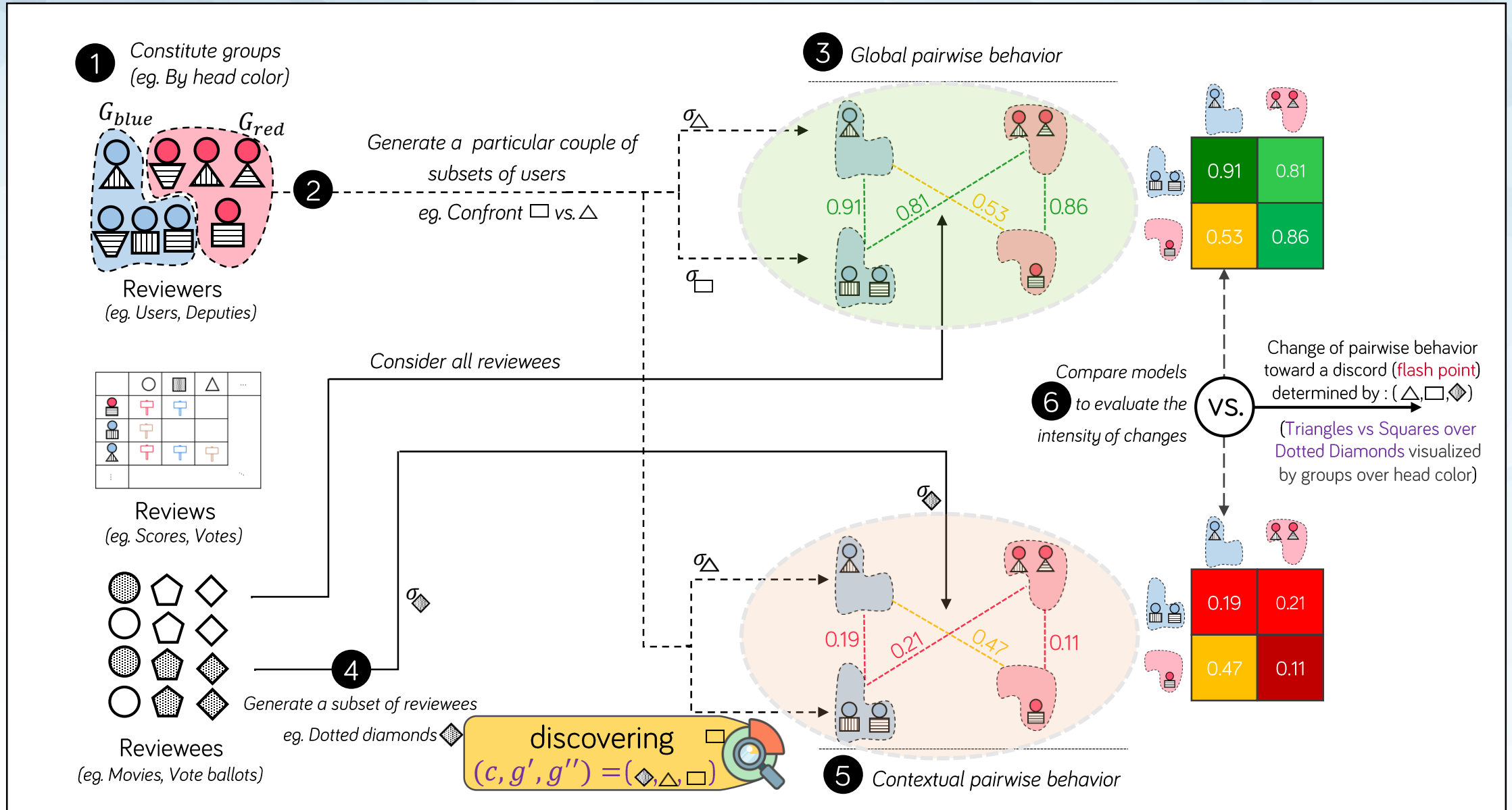
Change of pairwise behavior  
toward a discord (flash point)  
determined by : ( $\triangle$ ,  $\square$ ,  $\blacklozenge$ )  
(Triangles vs Squares over  
Dotted Diamonds visualized  
by groups over head color)

**5** Contextual pairwise behavior

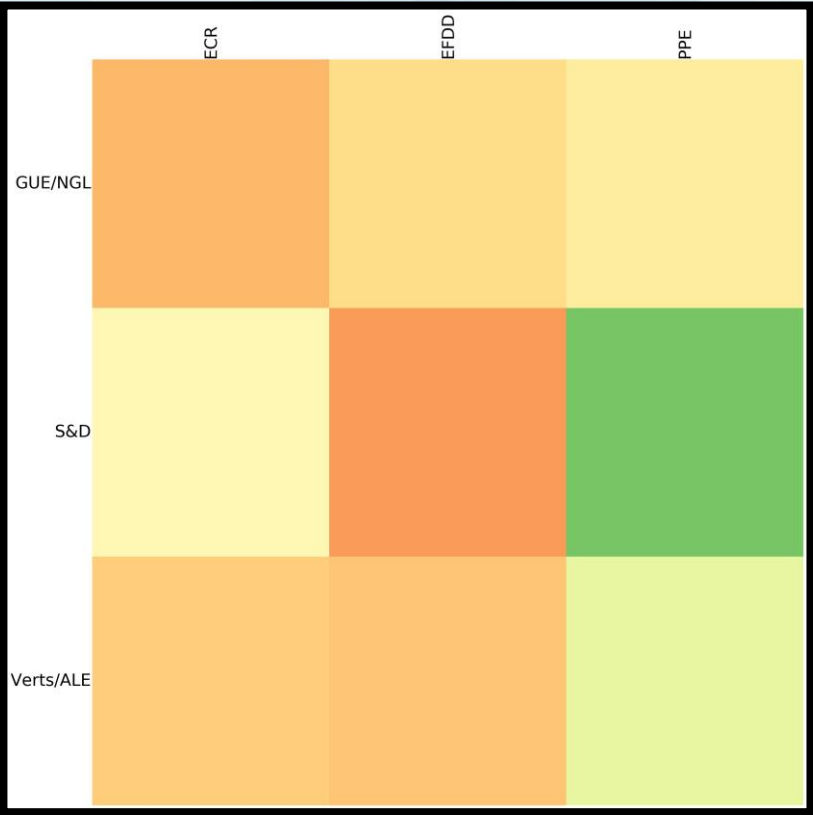


$\text{reviewer icon}$	$\text{reviewer icon}$
0.19	0.21
0.47	0.11

# DISCOVERING SIMILARITIES CHANGE ( DSC ) FRAMEWORK



### Usual pairwise behavior

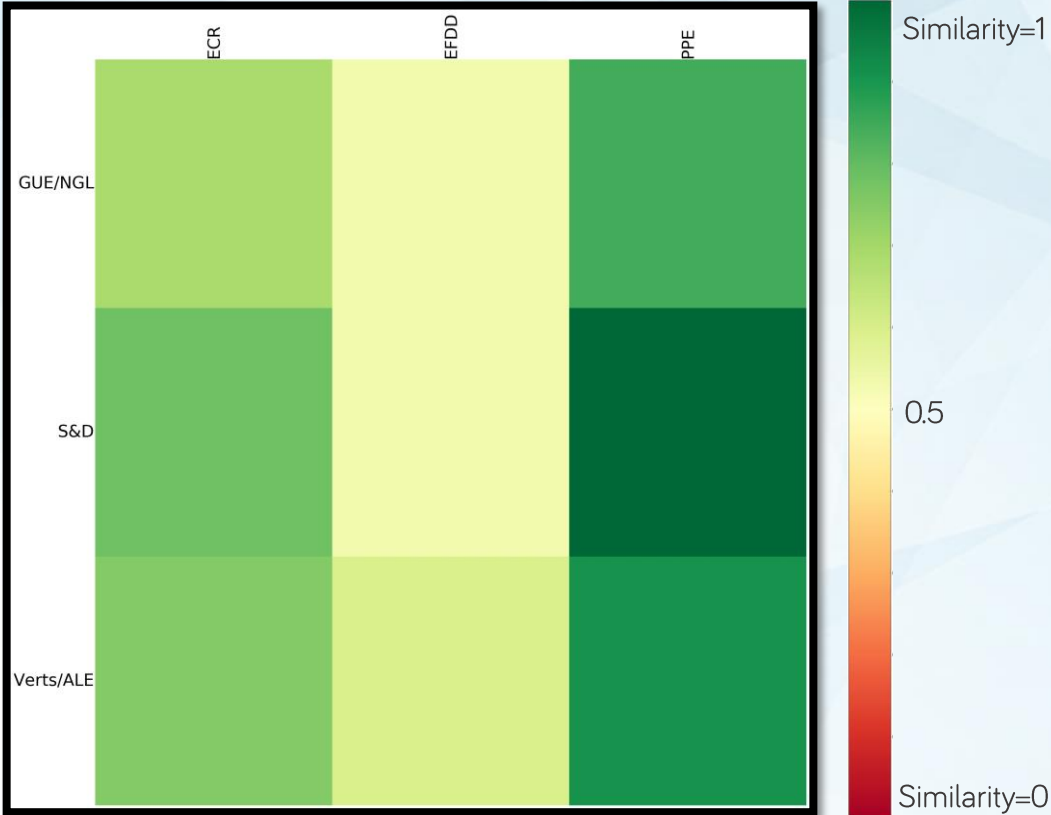


Toward **consent** between European political groups  
The pattern:



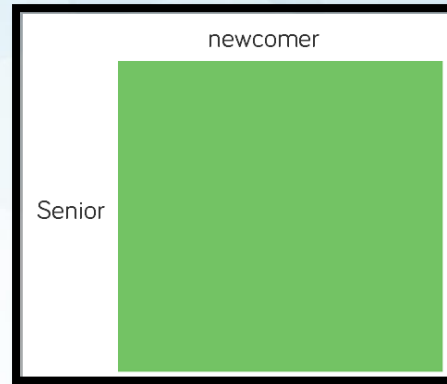
[7.40 European judicial conventions during Feb. - Nov. 2015 ,left wing ,right wing]

### Contextual pairwise behavior

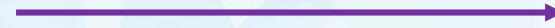




## Usual pairwise behavior

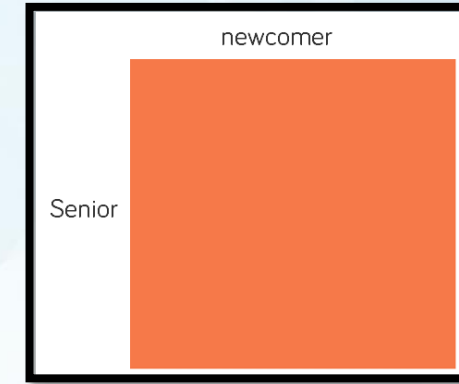


Toward **Dissent** between Yelp Users for the context:



[Professional Services, Shopping, In Oklahoma, Senior, Newcomer]

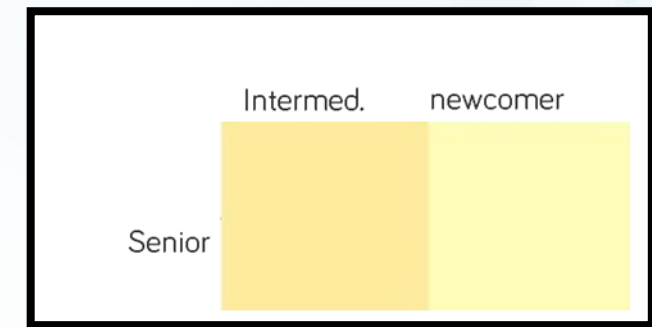
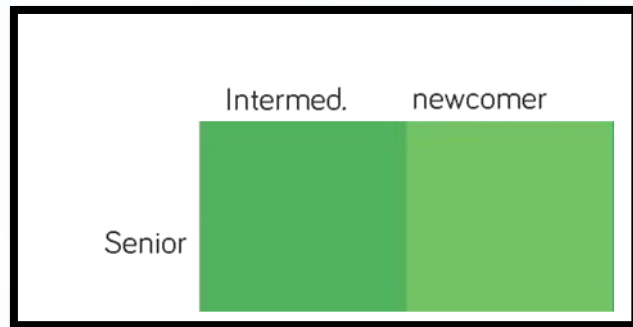
## Contextual pairwise behavior

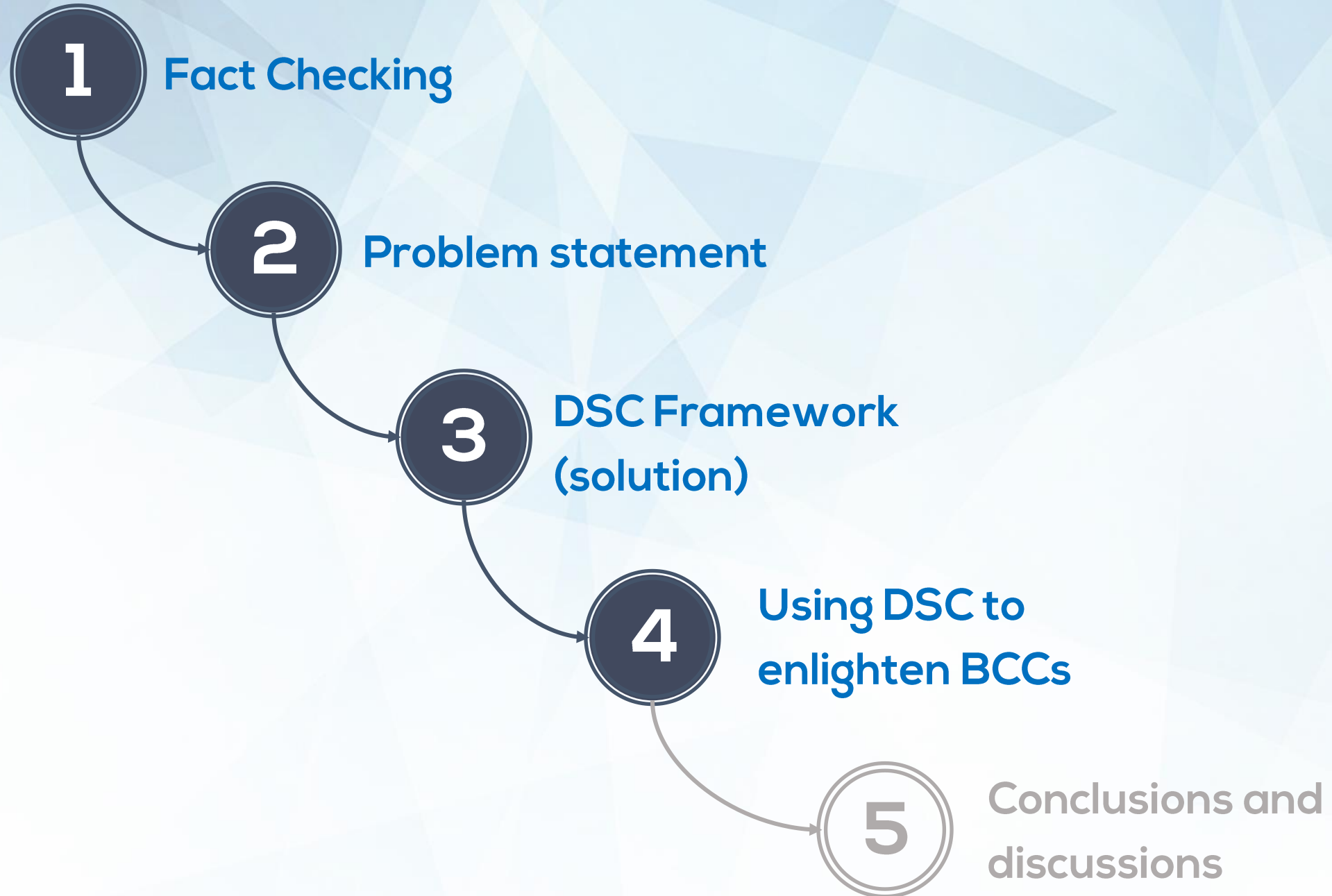


Toward **Dissent** between Yelp Users for the context:



[Medical Center, Doctors, In Wisconsin, {Senior}, {Intermed., Newcomer,}]







**Fact checking** is the act of **checking factual assertions** in non-fictional text in order to **determine the veracity and correctness of the factual statements** in the text.



We focus on specific type of claims namely **Behaviors Comparison Claims**.

- \* **Behaviors Comparison Claims** are statements that assert a similarity or a dissimilarity of behavior between individuals, groups, countries ...
- \* **Several Claims** can be transformed into **BCCs** thus allowing them to be **contextualized**.

**Example:** In the European Parliament, The deputy X votes practically the same as the deputy Y.

To evaluate **at what extent this claim is valid**, several questions pop in mind implying different ways to **enlighten such claims**:

- Is this claim valid in the general case (when considering all votes) ? *(do the math)*
- Contextualize the claim:
  - Considering that the two deputies are from two different parties, are the two respective parties similar in their voting behavior ?
  - Is it **valid for every context** (time period, topic of ballots) ?

**Claim 1** In the European parliament, French deputies vote following the votes recommendation given by their respective national parties

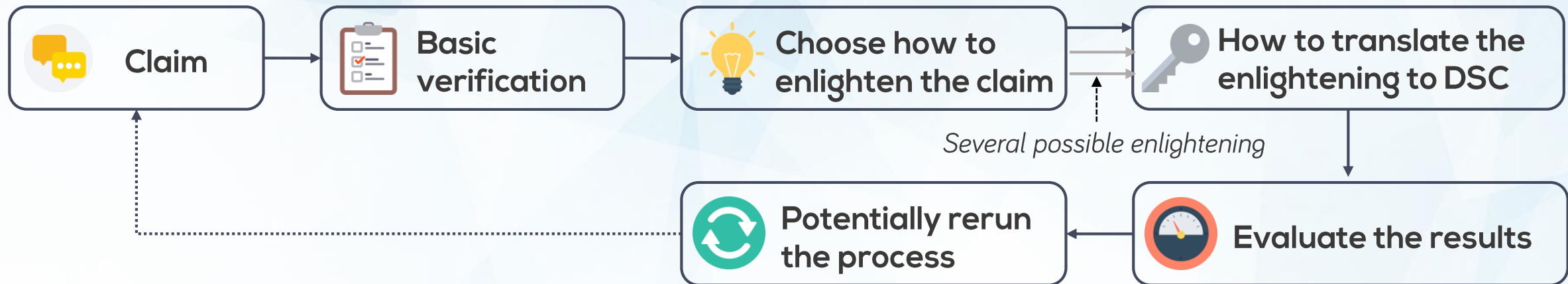
**Claim 2** There is no national position when it comes to votes in European political group.

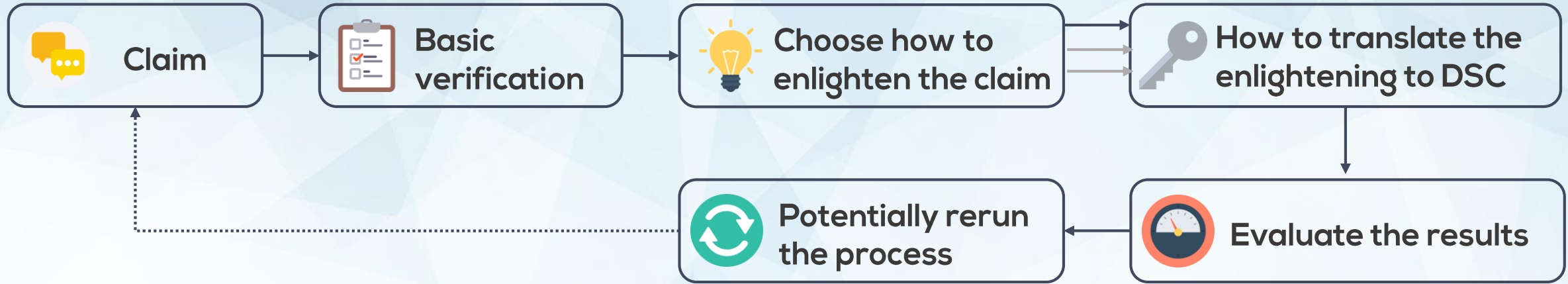
**Claim 3** Deputy D1 votes practically the same as a deputy D2 (Several possibilities by considering different dimensions of grouping ...)

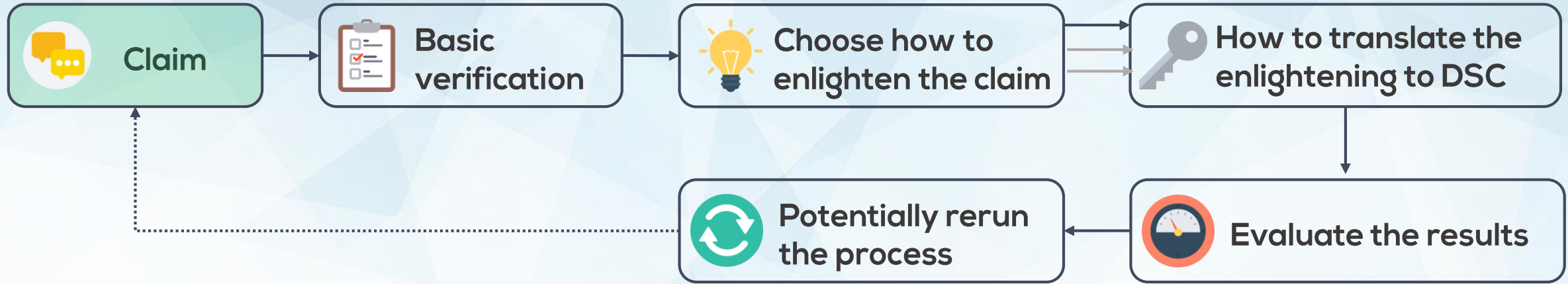
**Claim 4** The Topic X is a hotter than the Topic Y (w.r.t. all the parliament, some countries or some political groups ...)

**Claim 5** Deputy D1 changed his behavior after 2013 compared to its national party (the two political line diverge at some point after 2013 or for particular contexts)

What is the process of evaluating Behavioral Comparison Claims?

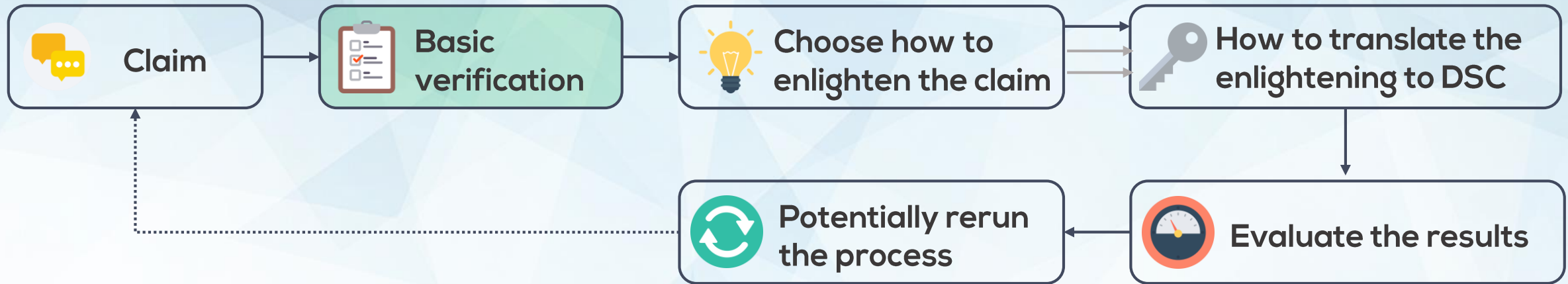






**Claim 1** In the European parliament, French deputies vote following the votes recommendation given by their respective national parties





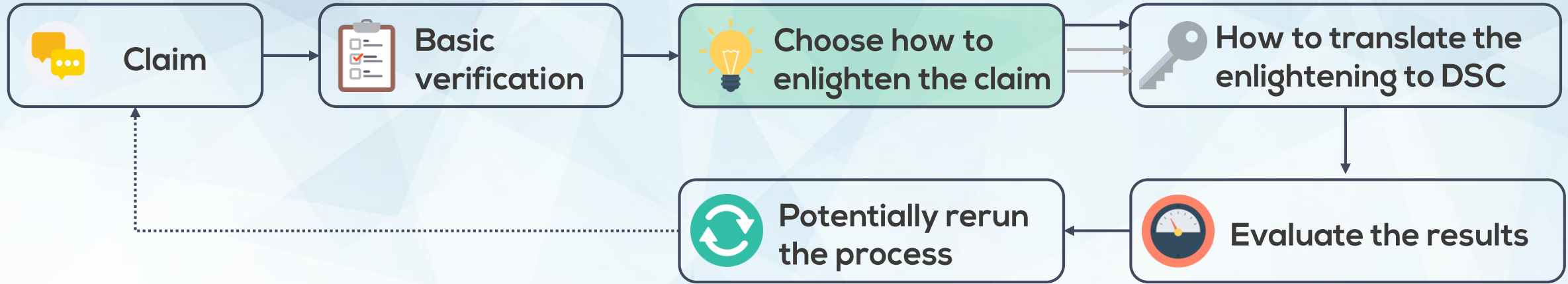
Basic  
Verification

For each French National Party and using all the roll call votes: evaluate the usual intra-cohesion

*\*To measure intra cohesion, the Agreement Index<sup>1</sup> can be used*

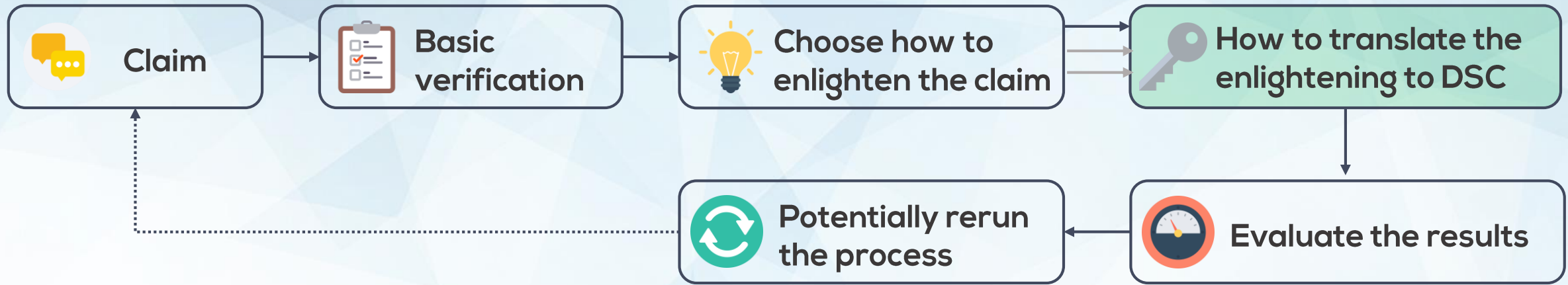


<sup>1</sup>Hix, S., Noury, A., & Roland, G. (2005). Power to the parties: cohesion and competition in the European Parliament, 1979–2001. *British Journal of Political Science*, 35(2), 209-234.



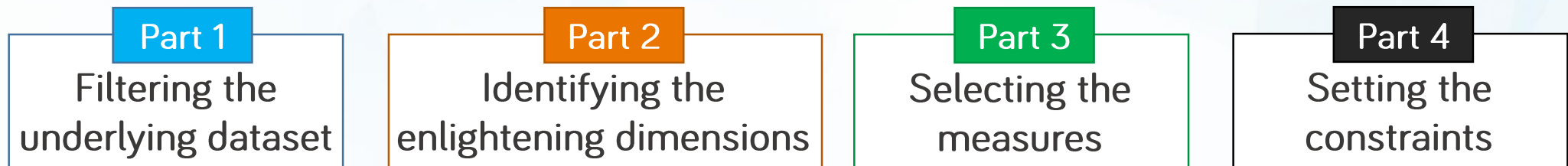
*Enlightening  
the claim*

Finding counter arguments: is there any particular context (identifying a subset of ballots) where a national party intra-cohesion decreases significantly?



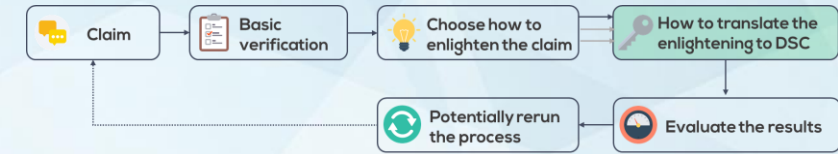
 1  
*Translating  
(Using DSC)*

Discovering Intra-cohesion changes: We are then interested in finding contexts (subsets of items) where the intra-cohesion measure for a national party reduces w.r.t. its usual intra-cohesion.



## 1 Translating (Using DSC)

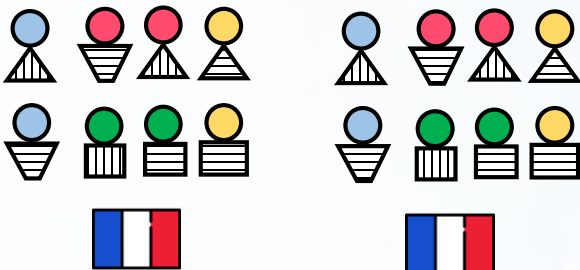
Discovering Intra-cohesion changes: We are then interested in finding contexts (subsets of items) where the intra-cohesion measure for a national party reduces w.r.t. its usual intra-cohesion.



## Part 1 Filtering the underlying dataset

1 What are the two subsets of individuals that I am going to confront in this study?

French Deputies against French deputies



2 On what subset of items am I going to build the referential behavior?

The usual behavior: then All ballots



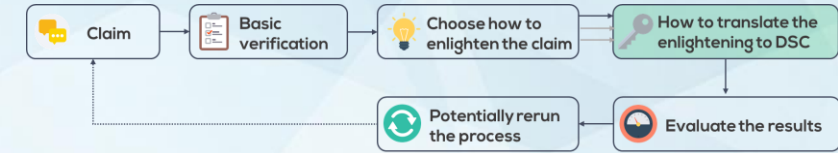
3 Do I have a prior knowledge on what contexts domains I want to explore?

No: consider then All ballots



1  
Translating  
(Using DSC)

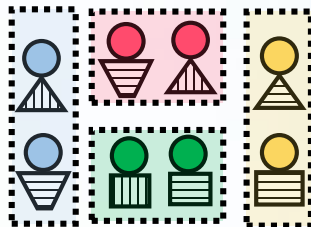
Discovering Intra-cohesion changes: We are then interested in finding contexts (subsets of items) where the intra-cohesion measure for a national party reduces w.r.t. its usual intra-cohesion.



## Part 2 Identifying the enlightening dimensions

4 What dimensions to consider when building groups of individuals?

National Parties



5 How a context is defined (by which ballots description attributes)?

The ballots themes

Items (Ballots) - <i>I</i>			Individuals (Deputies) - <i>U</i>			Outcome
Idsession	Date	Theme	Full name	National Party	Political Group	Vote
001	2017/03/17	1.10 Justice 2.10 Europe coop	Lavrilieux	LR	PPE	For
001	2017/03/17	1.10 Justice 2.10 Europe coop	Philippot	FN	ENF	Against
002	2017/04/11	3.10 Agriculture	Lavrilieux	LR	PPE	For
002	2017/04/11	3.10 Agriculture	Philippot	FN	ENF	For
002	2017/04/11	3.10 Agriculture	Arnatu	FN	ENF	For
003	2017/04/11	1.20 Security	Le Grip	LR	PPE	Abstain

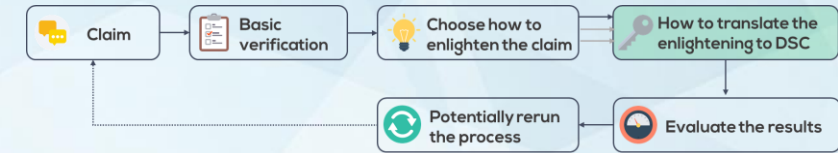
6 How a subset of individuals is defined (by which deputies description attributes)?

National Parties: + Confront only the same subsets (Intra groups behavior)

Items (Ballots) - <i>I</i>			Individuals (Deputies) - <i>U</i>			Outcome
Idsession	Date	Theme	Full name	National Party	Political Group	Vote
001	2017/03/17	1.10 Justice 2.10 Europe coop	Lavrilieux	LR	PPE	For
001	2017/03/17	1.10 Justice 2.10 Europe coop	Philippot	FN	ENF	Against
002	2017/04/11	3.10 Agriculture	Lavrilieux	LR	PPE	For
002	2017/04/11	3.10 Agriculture	Philippot	FN	ENF	For
002	2017/04/11	3.10 Agriculture	Arnatu	FN	ENF	For
003	2017/04/11	1.20 Security	Le Grip	LR	PPE	Abstain

## 1 Translating (Using DSC)

Discovering Intra-cohesion changes: We are then interested in finding contexts (subsets of items) where the intra-cohesion measure for a national party reduces w.r.t. its usual intra-cohesion.



## Part 3 Selecting the measures

7 What similarity measures between confronted subsets of individuals do you want to use?

*Agreement index*

8 Are you interested in strengthening similarities or weakening similarities?

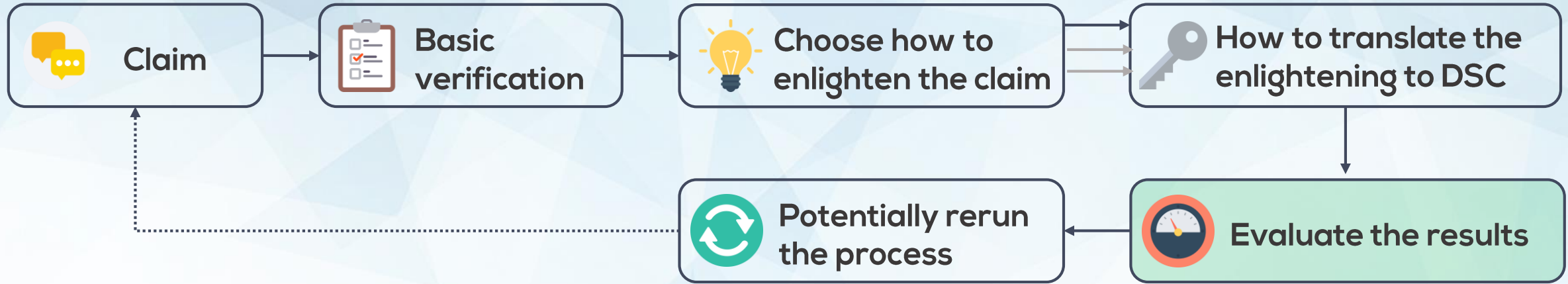
*Weakening of similarities: Decreases of intra cohesion*

## Part 4 Setting the constraints

9 Advanced constraints and other info:

- What are the minimum number of ballots over which a resulting pattern is viewed as significant? (#thres. Items= 10)
- What is the minimum number of deputies composing a group? (#thres. Items= 5)
- How much results do you want to display? (Top-K) (K= 25)

...



*Evaluating  
the results*

Results: DSC gives the set of the most significant patterns w.r.t. the Intensity of change of pairwise behavior.

# FACT CHECKING, PLAYING THE PROCESS



Evaluating the results

Results: DSC gives the set of the most significant patterns w.r.t. the Intensity of change of pairwise behavior.



Summary of patterns found

index	pattern	context	g1	g2	[subgroup(context)]	[subgroup(g1)]	[subgroup(g2)]	#reviews	quality	upperbound	items_details
0	[[3.70.12]], [[Front national]], [[Front national]]	[[3.70.12 Waste management, domestic waste, packaging, light industrial waste]]	[[Front national]]	[[Front national]]	44	23	23	948	0.33	0.41906889527347757	['Resource efficiency: moving towards a circular economy', 'Packaging and packaging waste: reduction of the consumption of lightweight plastic carrier bags']

Showing 1 to 1 of 10 entries

Reviews of pattern

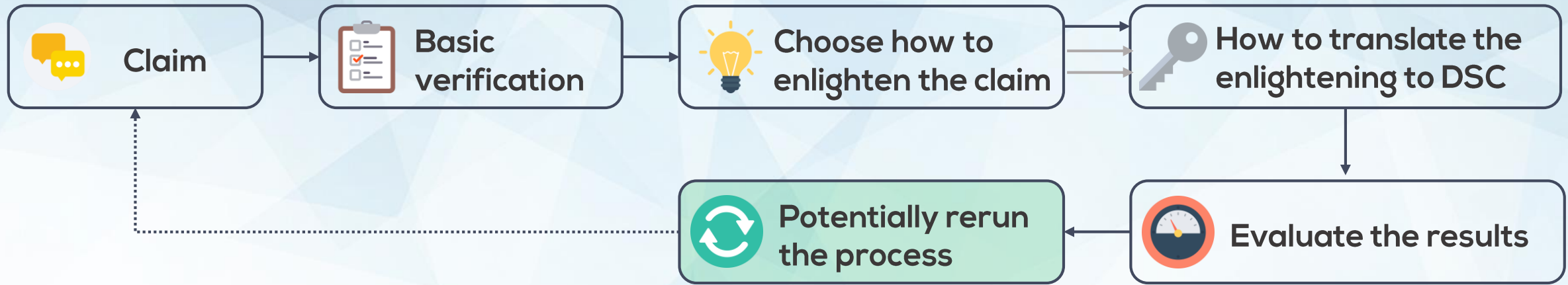
VOTEID	PROCEDURE_TITLE	VOTE_DATE	VOTE_DATE_DETAILED	PROCEDURE_SUBJECT	DOSSIERID	PROCEDURE_SUBTYPE	COMMITTEE	PROCEDURE_TYPE	EP_ID	NAME_FULL	NATIONAL_PARTY	GROUPE_ID	COUNTRY	AGEGROUP	GENDER	AGE	USER_VOTE
--------	-----------------	-----------	--------------------	-------------------	-----------	-------------------	-----------	----------------	-------	-----------	----------------	-----------	---------	----------	--------	-----	-----------

Referential Heatmap

Contextual Heatmap

For the context -Waste management-, The intra-cohesion of the French national party FN tends to decrease by 30%





Rerun the process

Depending on the claim, it may be interesting to enlighten it from another angle of view.

*\*We will give an example of this for the next claim.*

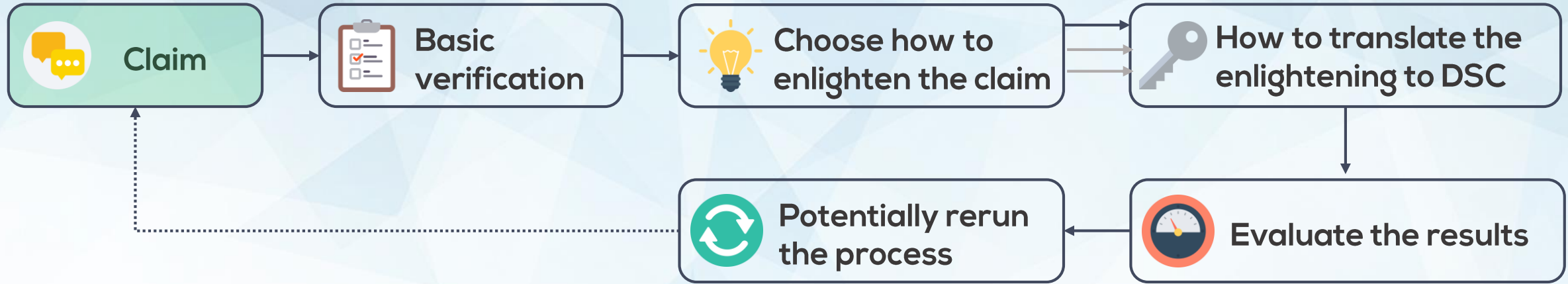
**Claim 1** In the European parliament, French deputies vote following the votes recommendation given by their respective national parties → *Valid in general case, but there is some particular contexts where deputies of certain national parties are divided*

**Claim 2** There is no national position when it comes to votes in European political group.

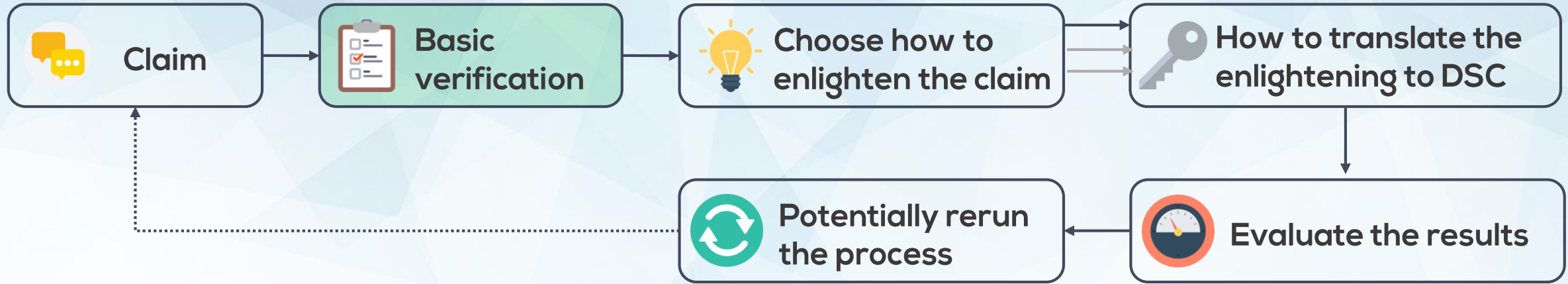
**Claim 3** Deputy D1 votes practically the same as a deputy D2 (Several possibilities by considering different dimensions of grouping ...)

**Claim 4** The Topic X is a hotter than the Topic Y (w.r.t. all the parliament, some countries or some political groups ...)

**Claim 5** Deputy D1 changed his behavior after 2013 compared to its national party (the two political line diverge at some point after 2013 or for particular contexts)



**Claim 2** There is no national position when it comes to votes in European political group (Select S&D for example).

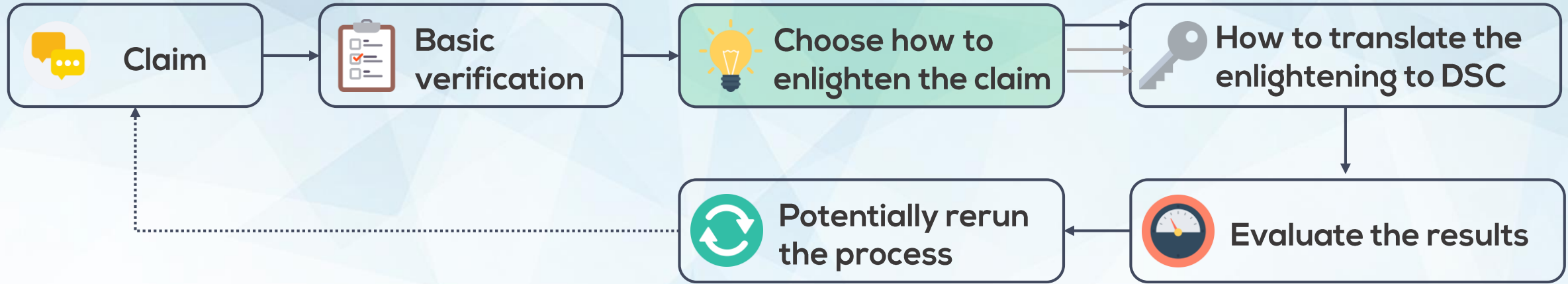


*Basic  
Verification*

For each European Political Group and using all the roll call votes: evaluate the usual intra-cohesion for a given political group (There are 7)



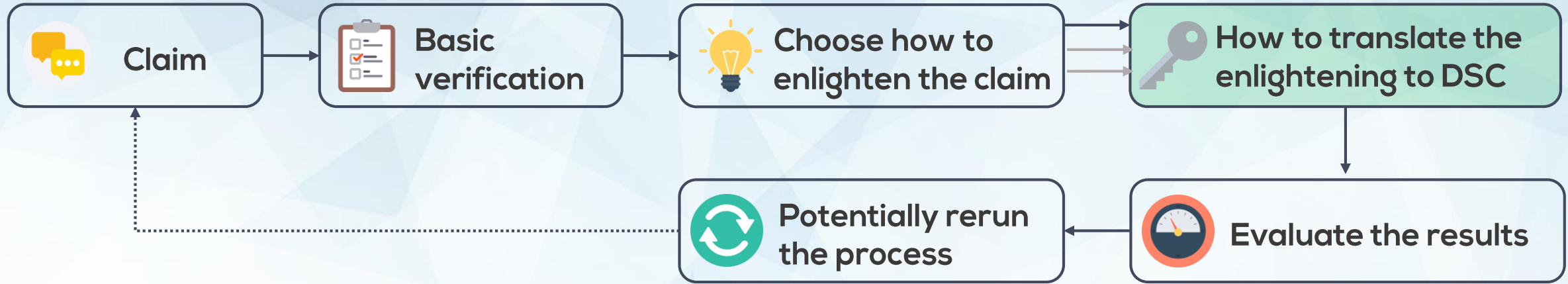
For a given European Political Group, confront deputies of each peer of countries and using all the roll call votes: evaluate the usual pairwise behavior.



Enlightening  
the claim

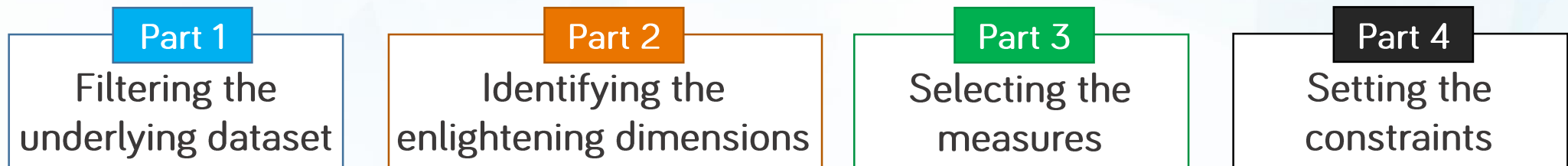
Finding counter arguments/or strengthening arguments: is there any particular context (identifying a subset of ballots) where for a given European group we have an important Weakening/Strengthening of the usual observed pairwise behavior\*

*\*similarity between the two confronted majorities of countries*



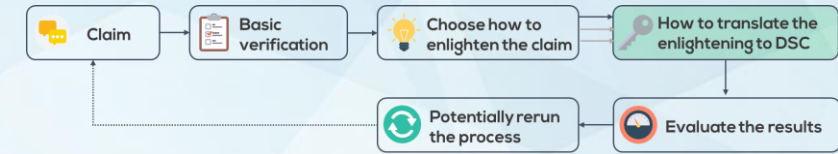
 **2**  
*Translating  
(Using DSC)*

Discovering change of pairwise agreement: We are then interested in finding contexts (subsets of items) where the pairwise behavior between two confronted countries of the same political group changes drastically (e.g.: toward discord) w.r.t. its usual maintained pairwise behavior.



## 2 Translating (Using DSC)

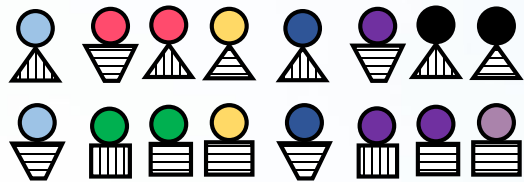
Discovering change of pairwise agreement: We are then interested in finding contexts (subsets of items) where the pairwise behavior between two confronted countries of the same political group changes drastically (e.g.: toward discord) w.r.t. its usual maintained pairwise behavior.



## Part 1 Filtering the underlying dataset

1 What are the two subsets of individuals that I am going to confront in this study?

All Deputies against All Deputies



2 On what subset of items am I going to build the referential behavior?

The usual behavior: then All ballots



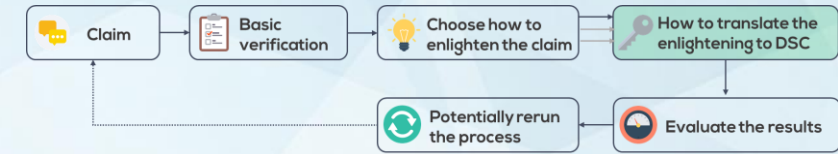
3 Do I have a prior knowledge on what contexts domains I want to explore?

No: consider then All ballots



**2**  
*Translating  
(Using DSC)*

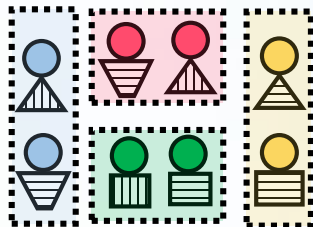
Discovering change of pairwise agreement: We are then interested in finding contexts (subsets of items) where the pairwise behavior between two confronted countries of the same political group changes drastically (e.g.: toward discord) w.r.t. its usual maintained pairwise behavior.



## Part 2 Identifying the enlightening dimensions

**4** What dimensions to consider when building groups of individuals?

Countries



**5** How a context is defined (by which ballots description attributes)?

The ballots themes & period of voting

Items (Ballots) - <i>I</i>			Individuals (Deputies) - <i>U</i>			Outcome
Idsession	Date	Theme	Full name	National Party	Political Group	Vote
001	2017/03/17	1:10 Justice 2:10 Europe coop	Lavrilieux	LR	PPE	For
001	2017/03/17	1:10 Justice 2:10 Europe coop	Philippot	FN	ENF	Against
002	2017/04/11	3:10 Agriculture	Lavrilieux	LR	PPE	For
002	2017/04/11	3:10 Agriculture	Philippot	FN	ENF	For
002	2017/04/11	3:10 Agriculture	Arnatu	FN	ENF	For
003	2017/04/11	1:20 Security	Le Grip	LR	PPE	Abstain

**6** How a subset of individuals is defined (by which deputies description attributes)?

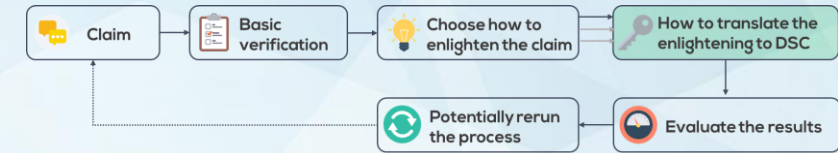
Countries:

Items (Ballots) - <i>I</i>			Individuals (Deputies) - <i>U</i>			Outcome
Idsession	Date	Theme	Full name	National Party	Political Group	Vote
001	2017/03/17	1:10 Justice 2:10 Europe coop	Lavrilieux	LR	PPE	For
001	2017/03/17	1:10 Justice 2:10 Europe coop	Philippot	FN	ENF	Against
002	2017/04/11	3:10 Agriculture	Lavrilieux	LR	PPE	For
002	2017/04/11	3:10 Agriculture	Philippot	FN	ENF	For
002	2017/04/11	3:10 Agriculture	Arnatu	FN	ENF	For
003	2017/04/11	1:20 Security	Le Grip	LR	PPE	Abstain



## 2 Translating (Using DSC)

Discovering change of pairwise agreement: We are then interested in finding contexts (subsets of items) where the pairwise behavior between two confronted countries of the same political group changes drastically (e.g.: toward discord) w.r.t. its usual maintained pairwise behavior.



## Part 3 Selecting the measures

7 What similarity measures between confronted subsets of individuals do you want to use?

Similarity between majority vote

8 Are you interested in strengthening similarities or weakening similarities?

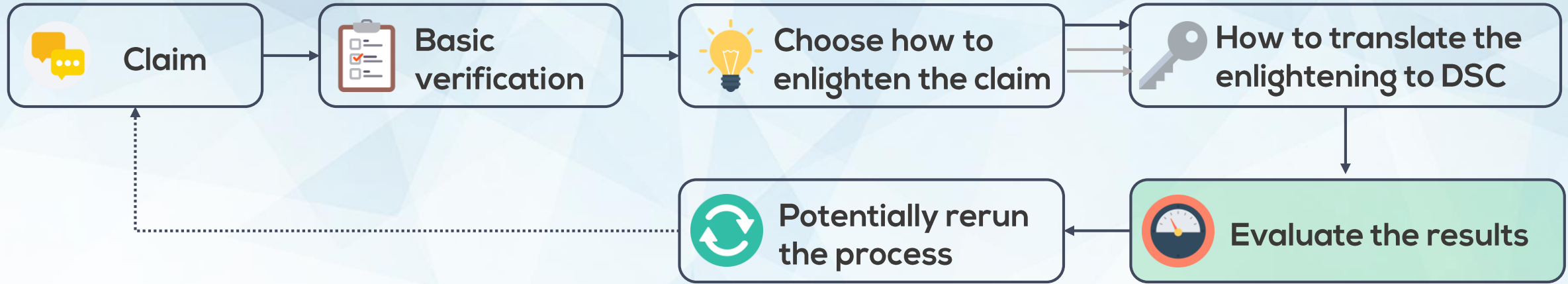
Weakening of similarities: Decreases of pairwise behavior

## Part 4 Setting the constraints

9 Advanced constraints and other info:

- What are the minimum number of ballots over which a resulting pattern is viewed as significant? (#thres. Items= 15)
- What is the minimum number of deputies composing a group? (#thres. Items= 20)
- How much results do you want to display? (Top-K) (K= 25)

...



*Evaluating  
the results*

Results: DSC gives the set of the most significant patterns w.r.t. the Intensity of change of pairwise behavior.

# FACT CHECKING, PLAYING THE PROCESS



Evaluating the results

Results: DSC gives the set of the most significant patterns w.r.t. the Intensity of change of pairwise behavior.



For the context

- (2 Internal Market, 3, 4.60 Consumers protection in general Between) [Oct 2015, June 2016]- , The pairwise behavior between the German and the Italian S&D deputies tends to decrease by 60%

### Summary of patterns found

index	pattern	context	g1	g2	[subgroup(context)]	[subgroup(g1)]	[subgroup(g2)]	#reviews	quality	upperbound	Items_details
0	[[2; 3; 4.60; [12.0, 20.0]; [Germany]; [Italy]]]	[[2 Internal market, single market; [Germany]; [Italy]; [12.0, 20.0]]]	[Germany]	[Italy]	20	27	32	1071	0.8	0.994732576985413	[Open internet access; Possible extension of geographical indication protection of the European Union to non-agricultural products; Casesins and caseinates intended for human consumption; Insurance distribution; Recast; Personal protective equipment; Appliances burning gaseous fuels; Cableway installations; Unfair trading practices in the food supply chain]

Showing 1 to 1 of 25 entries

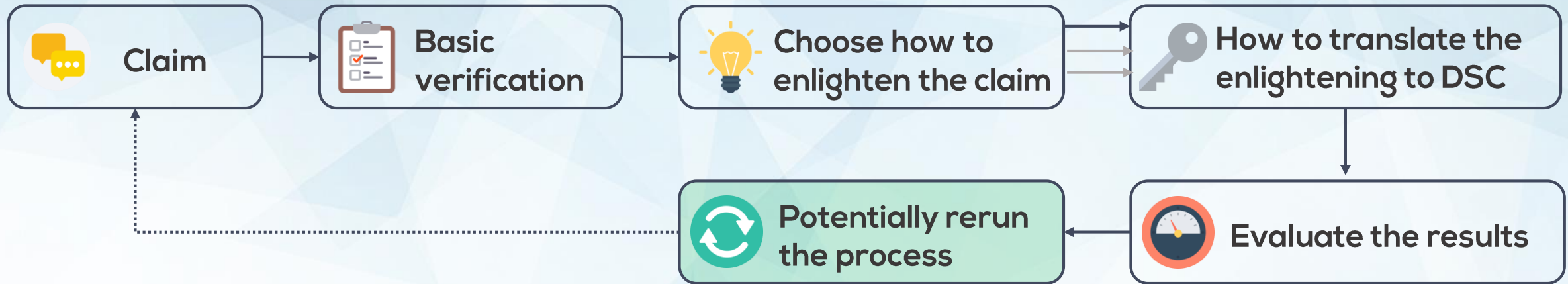
Previous 1 2 3 4 5 ... 25 Next

### Reviews of pattern

VOTEID	PROCEDURE_TITLE	VOTE_DATE	VOTE_DATE_DETAILED	PROCEDURE_SUBJECT	DOSSIERID	PROCEDURE_SUBTYPE	COMMITTEE	PROCEDURE_TYPE	EP_ID	NAME_FULL	NATIONAL_PARTY	GROUPE_ID	COUNTRY	AGEGROUP	GENDER	AGE	USER_VOTE
--------	-----------------	-----------	--------------------	-------------------	-----------	-------------------	-----------	----------------	-------	-----------	----------------	-----------	---------	----------	--------	-----	-----------

### Referential Heatmap

### Contextual Heatmap



*Rerun the process*

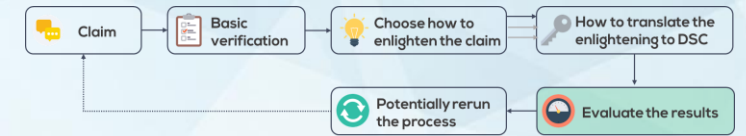
We seen that German and Italian S&D deputies are in disagreement considering the context shown before. Is it the a context that divide the two countries in general?

# FACT CHECKING, PLAYING THE PROCESS



Evaluating the results

Results: DSC gives the set of the most significant patterns w.r.t. the Intensity of change of pairwise behavior.



For the context  
- (2 Internal Market, 3) [Oct 2015, June 2016]- , The pairwise behavior between the German and the Italian ALL deputies tends to decrease by 51%

Summary of patterns found

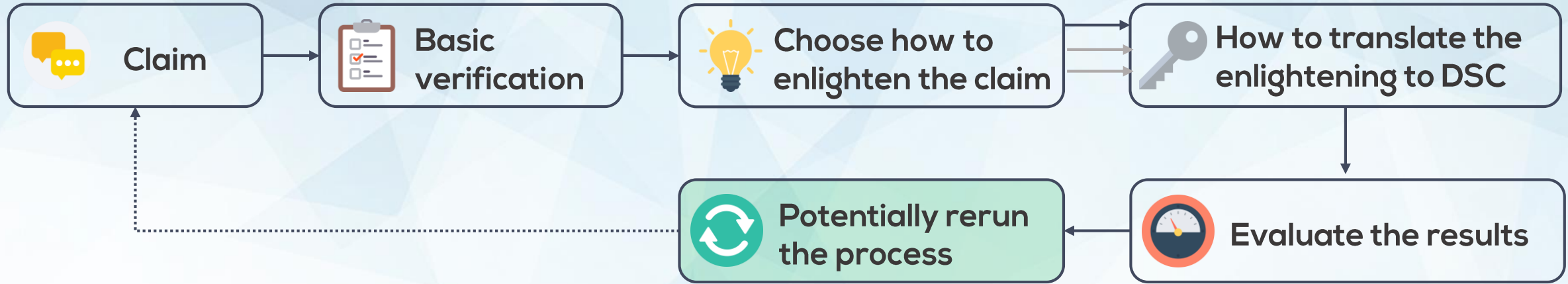
index	pattern	context	g1	g2	[subgroup(context)]	[subgroup(g1)]	[subgroup(g2)]	#reviews	quality	upperbound	items_details
0	[[2, 3], [12.0, 12.0], [Germany], [Italy]]	[[2 Internal market, single market, -], [12.0, 12.0]]	[Germany]	[Italy]	22	96	75	3501	0.92	0.9828207203561312	[Open internet access', 'Taxation: mandatory automatic exchange of information', 'Possible extension of geographical indication protection of the European Union to non-agricultural products', 'Caseins and caseinates intended for human consumption']

Showing 1 to 1 of 20 entries

Reviews of pattern

Referential Heatmap

Contextual Heatmap



And we can rerun crossing the national parties and so on ...

*Rerun the process*

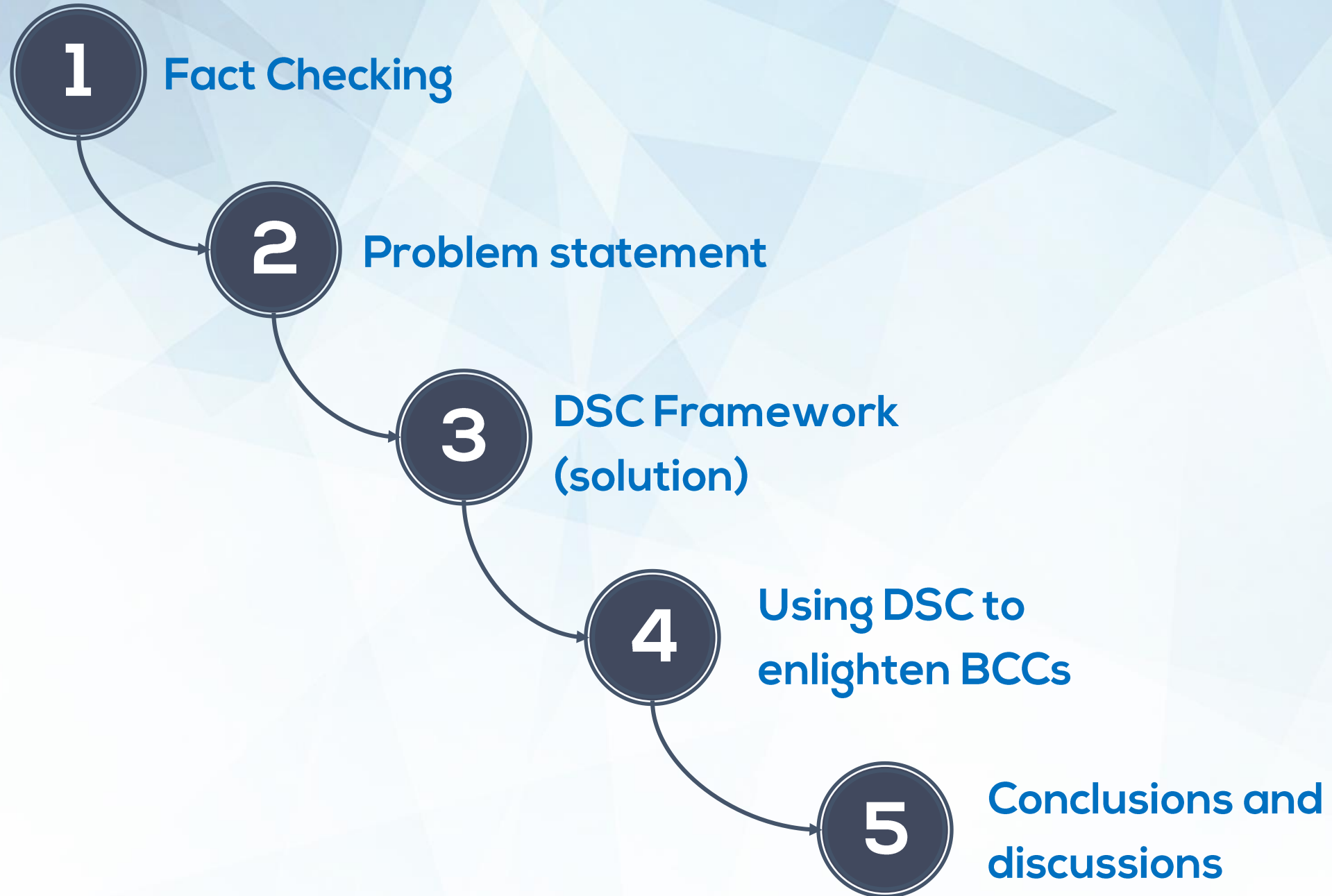
**Claim 1** In the European parliament, French deputies vote following the votes recommendation given by their respective national parties → *Valid in general case, but there is some particular contexts where deputies of certain national parties are divided*

**Claim 2** There is no national position when it comes to votes in European political group.

**Claim 3** Deputy D1 votes practically the same as a deputy D2 (Several possibilities by considering different dimensions of grouping ...)

**Claim 4** The Topic X is a hotter than the Topic Y (w.r.t. all the parliament, some countries or some political groups ...)

**Claim 5** Deputy D1 changed his behavior after 2013 compared to its national party (the two political line diverge at some point after 2013 or for particular contexts)





## Conclusion



- ◆ Definition of a novel problem :  
discovering exceptional pairwise behavior
- ◆ A little primer on evaluating/enlightening  
BCCs (Behaviors Comparison Claims)  
using DSC

## Conclusion



- ◆ Definition of a novel problem : discovering exceptional pairwise behavior
- ◆ A little primer on evaluating/enlightening BCCs (Behaviors Comparison Claims) using DSC



## Perspective

- ◆ Providing adapted instant mining and interactive mining algorithms.
- ◆ Studying the behavior of groups of individuals (deputies) through time.
- ◆ ContentCheck: Design a set of tools for fact checking/lead finding

THANKS

FOR YOUR TIME



QUESTIONS

Contact : [adnene.belfodil@insa-lyon.fr](mailto:adnene.belfodil@insa-lyon.fr)

Materials: <https://github.com/Adnene93/DiscoveringSimilarityChanges>

Feel free to ask any question you have :-)



# BIBLIOGRAPHY

THANKS  
FOR YOUR TIME



- ⑥ Belfodil, A., Cazalens, S., Lamarre, P., & Plantevit, M. (2017, September). Flash points: Discovering exceptional pairwise behaviors in vote or rating data. In ECML/PKDD.
- ⑥ Duivesteijn, W., Feelders, A. J., & Knobbe, A. (2016). Exceptional model mining. *Data Mining and Knowledge Discovery*, 30(1), 47-98.
- ⑥ Ganter, B., & Kuznetsov, S. (2001). Pattern structures and their projections. *Conceptual Structures: Broadening the Base*, 129-142.
- ⑥ de Sá, C. R., Duivesteijn, W., Soares, C., & Knobbe, A. (2016, October). Exceptional Preferences Mining. In *International Conference on Discovery Science* (pp. 3-18). Springer International Publishing.
- ⑥ G. Bosc, J. Golebiowski, M. Bensa, C. Robardet, M. Plantevit, J. Boulicaut, and M. Kaytoue. Local subgroup discovery for eliciting and understanding new structure-odor relationships. In *DS*, pages 19-34, 2016.
- ⑥ Amer-Yahia, S., Kleisarchaki, S., Kolloju, N. K., Lakshmanan, L. V., & Zamar, R. H. (2017, April). Exploring Rated Datasets with Rating Maps. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 1411-1419). International World Wide Web Conferences Steering Committee.

**Contact :** [adnene.belfodil@insa-lyon.fr](mailto:adnene.belfodil@insa-lyon.fr)

**Materials:** <https://github.com/Adnene93/DiscoveringSimilarityChanges>

*Feel free to ask any question you have :-)*