Further Strategy beyond Checklist Survey

National Institute of Informatics

Masashi KAWAI

CC BY

JPCOAR

Japan Consortium for Open Access Repository

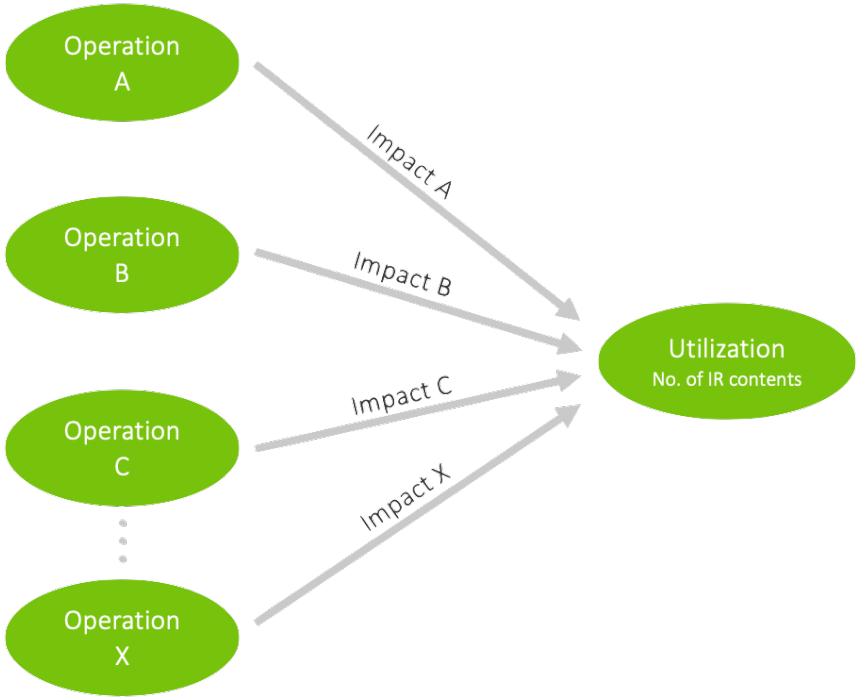
Background

- ✓ After conducting the checklist survey, data of many institutions about IR operations will be available.
- ✓ For instance, we will be able to obtain data about an operation relating to metadata management from a question in the checklist such as "Do you keep track of the metadata versions?"
- ✓ What we need to do after obtaining data from the checklist is identifying operations that have an impact on IR utilization
- ✓ It is because once those operations are identified, it will be able to improve IR utilization which is the main purpose of having an IR by spreading them among the IR community

Objective

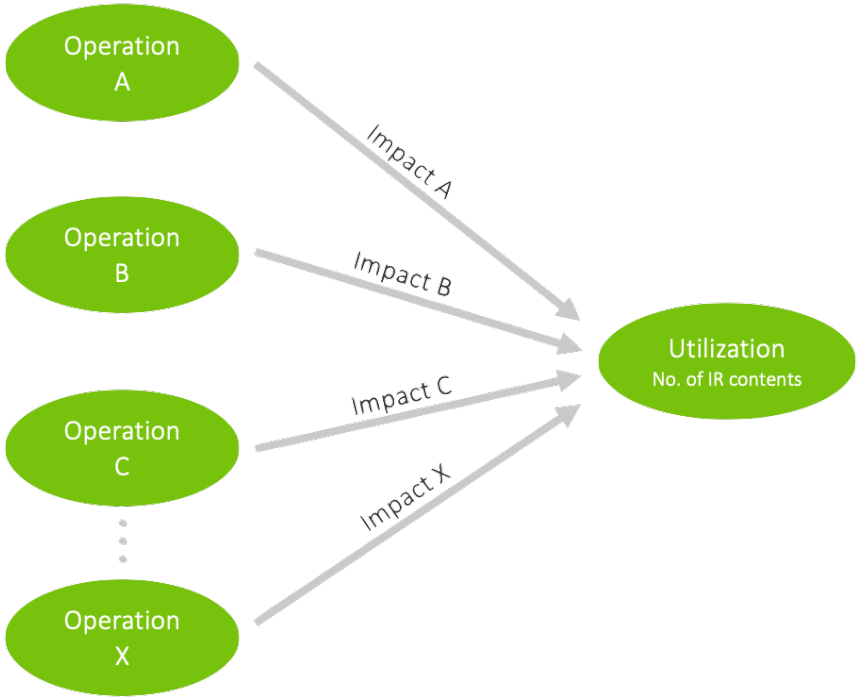
- ✓ The objective of the presentation is…
  : Identifying operations that have an impact on IR utilization
- ✓ In other words, the objective is…
  : Calculating each operation's impact on IR utilization

Image of Objective



Calculate impact

Image of Objective



Calculate impact → How ?

JPCOAR
Japan Consortium for Open Access Repository

Research Design

Image of Data Construction

1.Construct a table based on the responses to the checklist

	Operation A	Operation B	Operation C	...	Operation X	Utilization No. of IR contents
Instituion 1	Yes	No	Yes		No	1000
Instituion 2	Yes	No	Yes		No	2000
Instituion 3	Yes	No	No		Yes	500000
⋮						
Instituion Y	Yes	No	No		Yes	600000

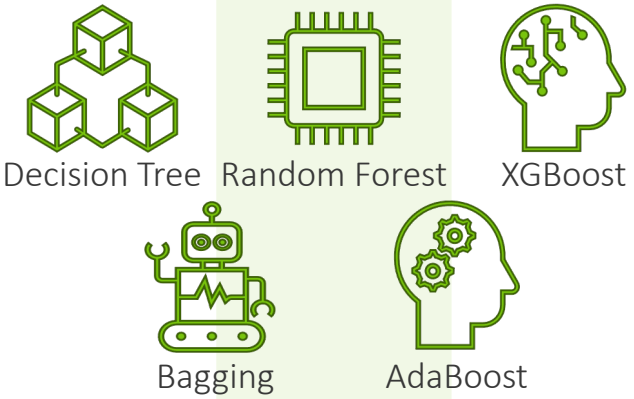
Note: Data is artificial

2.Change value type

	Operation A	Operation B	Operation C	...	Operation X	Utilization No. of IR contents
Instituion 1	1	0	1		0	1000
Instituion 2	1	0	1		0	2000
Instituion 3	1	0	0		1	500000
⋮						
Instituion Y	1	0	0		1	600000

Image of Model Construction

1.Select an algorithm



2. Generate a model based on the selected algorithm and the constructed table data

	Operation A	Operation B	Operation C	...	Operation X	Utilization No. of IR contents
Instituion 1	1	0	1		0	1000
Instituion 2	1	0	1		0	2000
Instituion 3	1	0			1	500000
⋮						
Instituion Y	1	0	0		1	600000

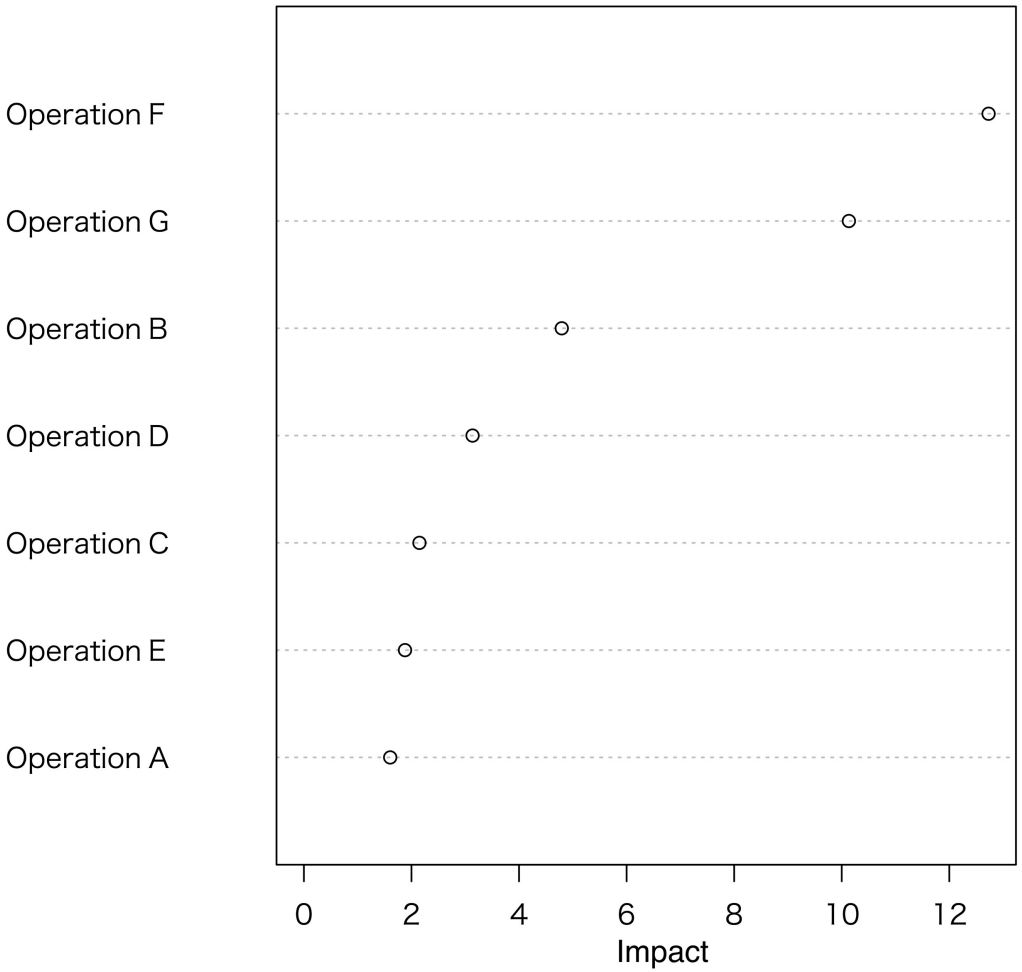
3. Calculate an impact

JPCOAR
Japan Consortium for Open Access Repository

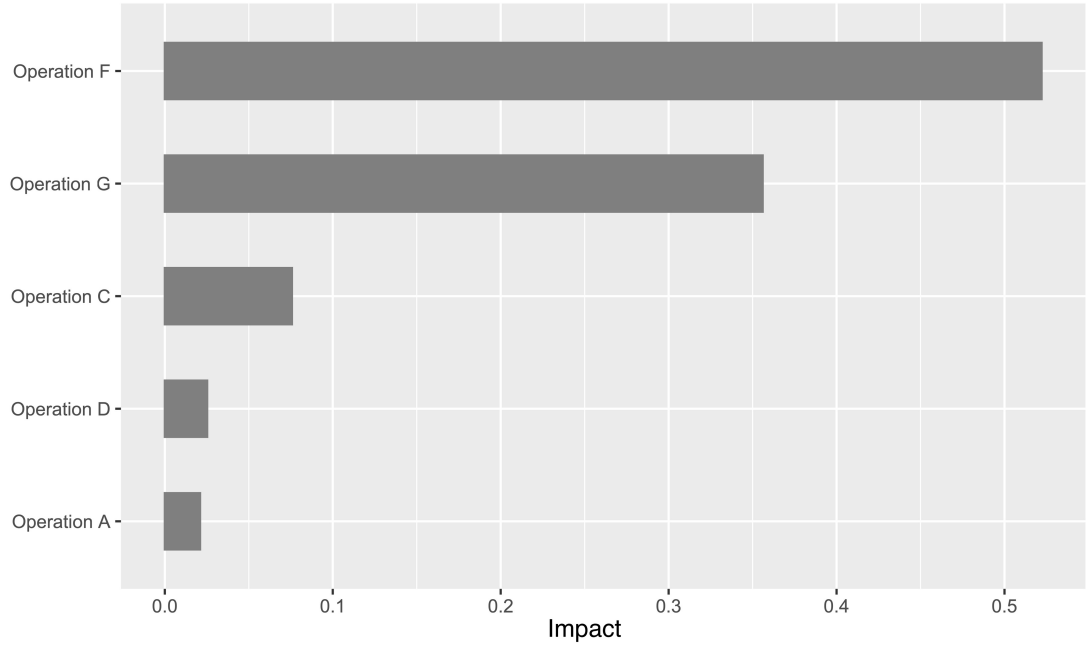
Analysis Results

JPCOAR
Japan Consortium for Open Access Repository

Random Forest



XGBoost


Operation F
Operation G
Operation C
Operation D
Operation A
0.0
0.1
0.2
0.3
0.4
0.5
Impact

Conclusion

Summary

- ✓ The checklist can be used to identify operations that have an impact on IR utilization
- ✓ Each operation's impact can be calculated through data and model construction
- ✓ Once each operation's impact is calculated, it will be able to improve IR utilization by spreading them among the IR community

Thanks for listening

m-Kawai@nii.ac.jp

JPCOAR

Japan Consortium for Open Access Repository