

Instructions for the demo application ALSowExt
in the following paper;

“Efficient extraction of related literature utilizing citation
networks by Academic-landscape/Future-prediction system;
Case study of enzyme-related literature selection”

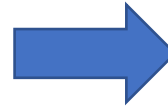
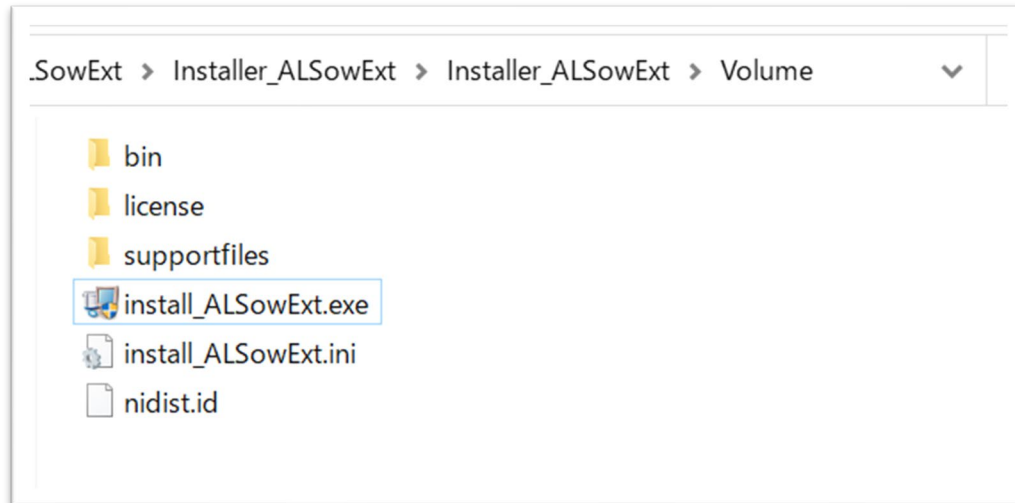
Hiroko Inoura and Nozomi Nagano

28th April 2023

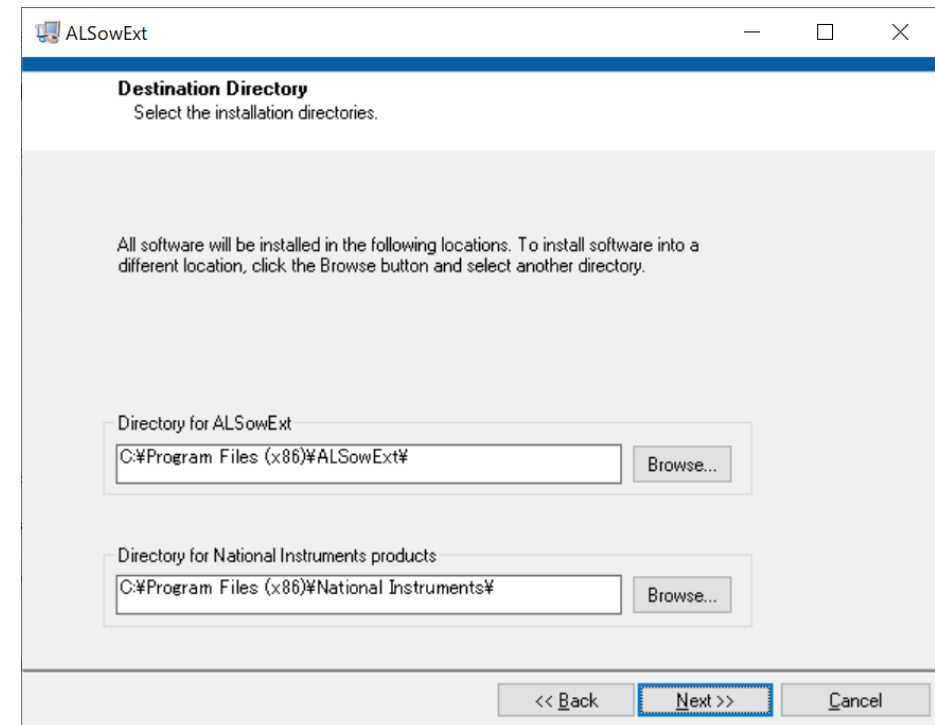
Installation of ALSowExt (for Windows 10)

By unzipping the file, installer_ALSowExt folder will be created. install_ALSowExt.exe file will be In the Volume folder.

Please execute it as an administrator.

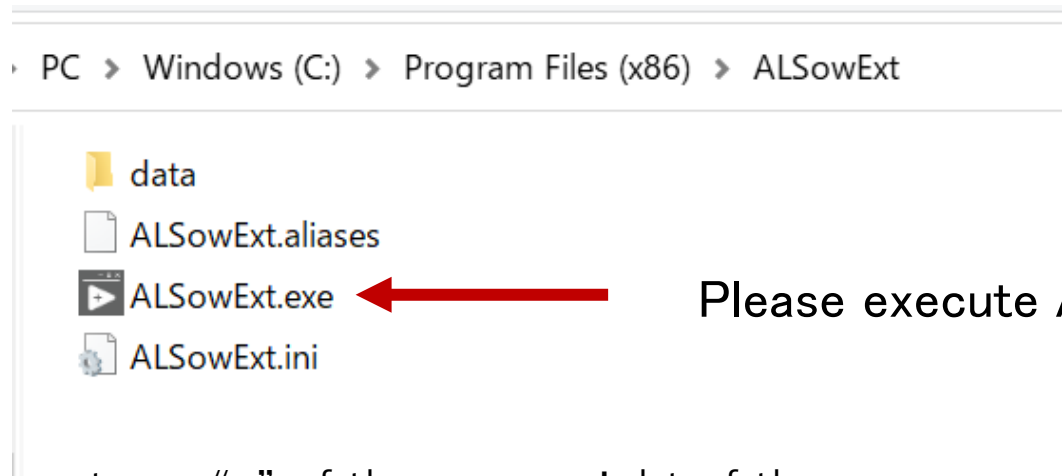


By executing the exe file, installer will appear. Please click the Next button to continue it..



Installation of ALSowExt (for Windows 10)

Finally, the ALSowExt folder will be created in the 'Program Files (x86)' folder.

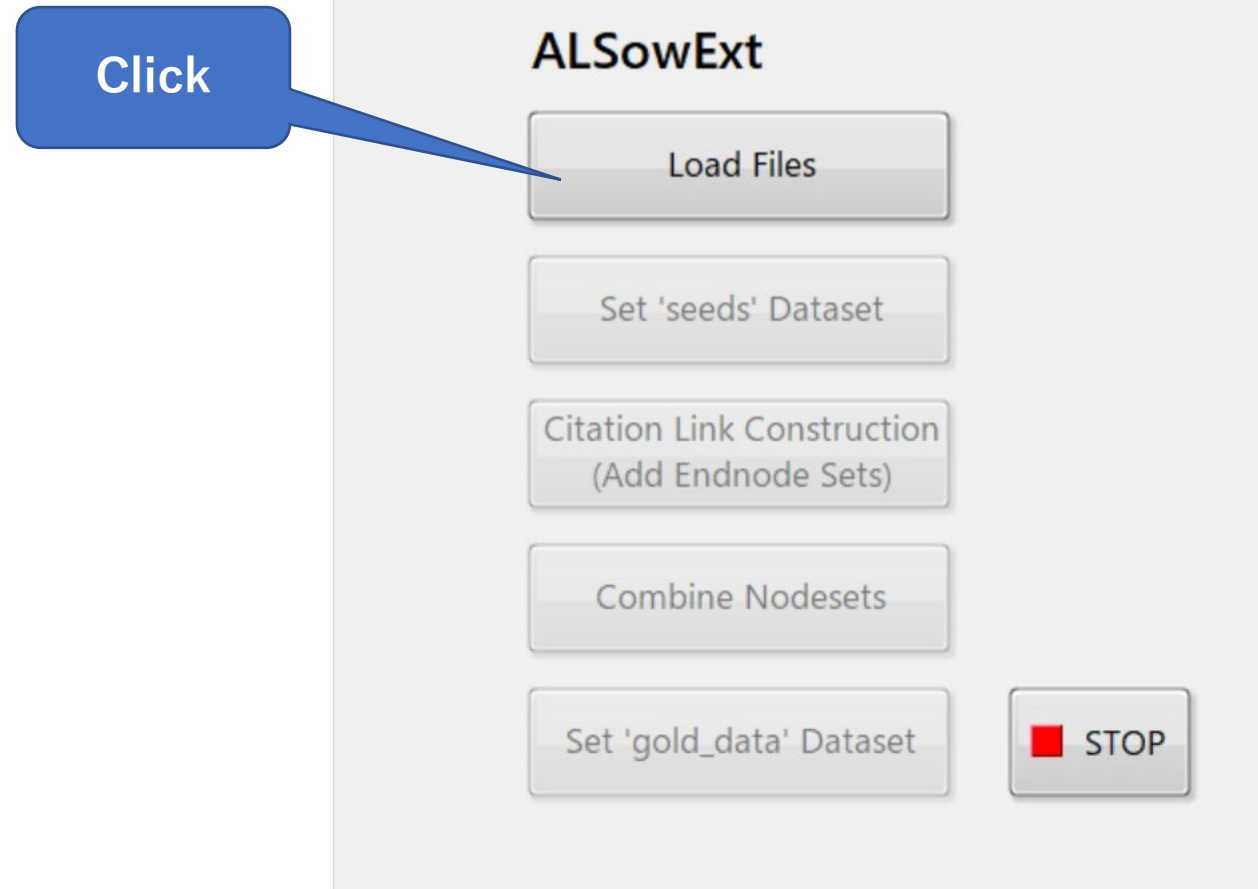


Please execute ALSowExt.exe to start it.

Warning: please do not use “x” of the upper-right of the window to close sub-windows. Please click the “back” (or “OK”, etc.) buttons to go back to the upper-level window.

Please move the ‘data’ folder to your Documents folder before executing ALSowExt.

Start of ALSowExt



Loading files

The image shows a 'load files' dialog box with two sections: 'FACET DATASET' and 'PAIR DATASET'. The 'FACET DATASET' section has a text field containing 'ALSowExt¥data¥a39818_facetall+D00197_EzCatRef-PDBsumKey.txt' and a folder icon button. The 'PAIR DATASET' section has a text field containing 'C:¥Users¥...xt¥data¥mission.pairs_a39818.tsv' and a folder icon button. At the bottom of the dialog are 'Go To Next' and 'Exit' buttons. To the right, a file explorer window shows a list of files. Two red arrows point from the file explorer to the folder icon buttons in the dialog. The first arrow points to the 'FACET DATASET' folder icon, and the second arrow points to the 'PAIR DATASET' folder icon. Both arrows have a red checkmark at their tails. The file explorer window shows a list of files, including 'a39565_facetall+S00021_EzCatRef-PDBsumKey.txt', 'a39576_facetall+D00165_EzCatRef-PDBsumKey.txt', 'a39577_facetall+S00225_EzCatRef-PDBsumKey.txt', 'a39589_facetall+D00471_EzCatRef-PDBsumKey.txt', 'a39593_facetall+T00074_EzCatRef-PDBsumKey.txt', 'a39818_facetall+D00197_EzCatRef-PDBsumKey.txt', 'mission.pairs_a39565.tsv', 'mission.pairs_a39576.tsv', 'mission.pairs_a39577.tsv', 'mission.pairs_a39589.tsv', 'mission.pairs_a39593.tsv', and 'mission.pairs_a39818.tsv'. The file 'a39818_facetall+D00197_EzCatRef-PDBsumKey.txt' and 'mission.pairs_a39818.tsv' are highlighted in yellow.

FACET DATASET ALSowExt¥data¥a39818_facetall+D00197_EzCatRef-PDBsumKey.txt

PAIR DATASET C:¥Users¥...xt¥data¥mission.pairs_a39818.tsv

Go To Next Exit

Please load the pair for the protein.
Relations of data detailed in the next slide.

Click to go to next step.
The window will close.

Use this button
to go back to
the top-level

Comparison of ALFP data and sample data file in this application

Datasets (EzCatDB ID)	ALFP analyzed data described in the paper	Data files included in this application
Lysozyme (S00021)	a39565/mission.facet.all.tsv	a39565_facetall+S00021_EzCatRef-PDBsumKey.txt
	a39565/mission.pairs.tsv	mission.pairs_a39565.tsv
a-Amylase (D00165)	a39576/mission.facet.all.tsv	a39576_facetall+D00165_EzCatRef-PDBsumKey.txt
	a39576/mission.pairs.tsv	mission.pairs_a39576.tsv
Triosephosphate isomerase (TIM) (S00225)	a39577/mission.facet.all.tsv	a39577_facetall+S00225_EzCatRef-PDBsumKey.txt
	a39577/mission.pairs.tsv	mission.pairs_a39577.tsv
HIV-1 retropepsin (D00471)	a39589/mission.facet.all.tsv	a39589_facetall+D00471_EzCatRef-PDBsumKey.txt
	a39589/mission.pairs.tsv	mission.pairs_a39589.tsv
Thrombin (T00074)	a39593/mission.facet.all.tsv	a39593_facetall+T00074_EzCatRef-PDBsumKey.txt
	a39593/mission.pairs.tsv	mission.pairs_a39593.tsv
Trypsin (D00197)	a39818/mission.facet.all.tsv	a39818_facetall+D00197_EzCatRef-PDBsumKey.txt
	a39818/mission.pairs.tsv	mission.pairs_a39818.tsv

Gold data and seed data were combined with facet.all data described in the paper, to make facetall+EzCatRef-PDBsumKey files.

Setting Seeds of Dataset for the further step

Click

Click

Click to go back to the top-level

Array index

scroll

STOP

SET seeds DATASET

BACK

seeds

_N 0 2693 3872 2637

ALSOwExt

File Edit View Operate Window Help

Load Files

Set 'seeds' Dataset

Citation Link Construction (Add Endnode Sets)

Combine Nodesets

Set 'gold_data' Dataset

set data

set dataset

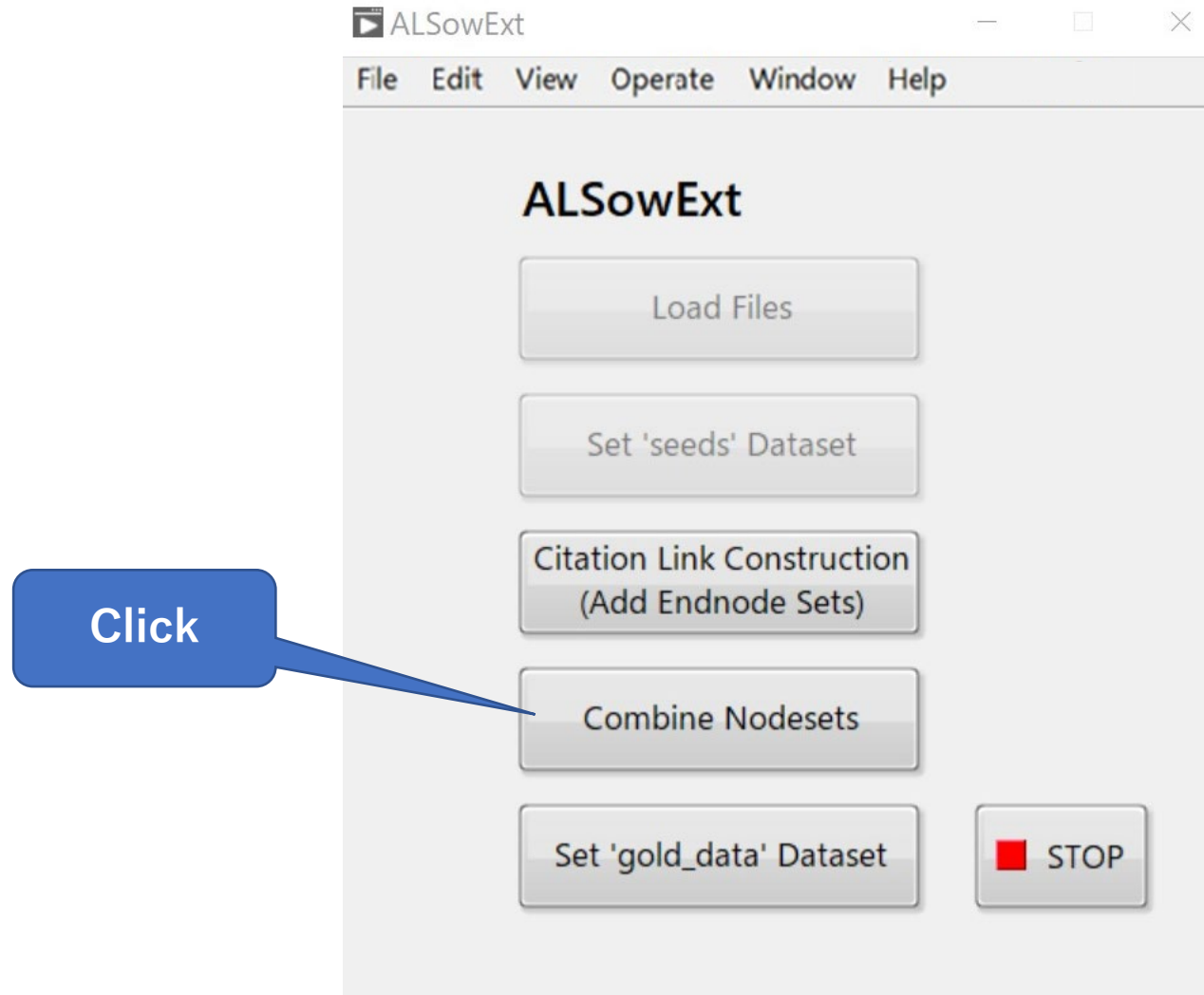
Please use scroll bar to explore the array, or increment/decrement button to change the index.

_N: record ID in the ALFP system

Citation Link Construction



Combining Node set



combine nodeset

1 To select a set, click the circle

2 Click the other set to combine

3 Select operant

4 Click to combine

scroll

Node sets/Endnode of Citation

SF

SR

OR (U)

COMBINE

Combine List

RESET Combine LIST

Count Endnode

Add to Node sets

BACK

Combined node set(s)	No. of endnode	ID(N)
	0	0 0 0 0 0
	0	0 0 0 0 0

combine nodeset

Node sets/Endnode of Citation Link

SF

SR

AND (\cap)

Node sets/Endnode of Citation Link

SF

SR

COMBINE

Combine List

[SF \cup SR]

[SF \cap SR]

RESET Combine LIST

Count Endnode

Add to Node sets

BACK

Combined node set(s)	No. of endnode	ID(_N)
	0	0 0 0 0 0
	0	0 0 0 0 0

5

Click to count end-nodes

Node sets/Endnode of Citation Link

SF ☐

SR ☐

☐

AND (\cap)

Node sets/Endnode of Citation Link

SF ☐

SR ☐

☐

COMBINE

RESET Combine LIST

Count Endnode

Add to Node sets

☐ BACK

Combine List

[SF \cup SR]
[SF \cap SR]

No. of end-nodes

Combined node set(s)

No. of endnode

ID(_N)

[SF \cup SR]

457

0

6

13

21

25

26

[SF \cap SR]

108

0

73

122

125

268

308

ALFP system IDs
in 1D Array

scroll

combine nodeset

Node sets/Endnode of Citation Link

SR

[SF \cup SR]

[SF \cap SR]

AND (\cap)

Node sets/Endnode of Citation Link

SF

SR

[SF \cup SR]

COMBINE

Combine List

[SF \cup SR]

[SF \cap SR]

RESET Combine LIST

Count Endnode

Add to Node sets

BACK

Combined node set(s)	No. of endnode	ID(_N)
[SF \cup SR]	457	6 13 21 25 26
[SF \cap SR]	108	73 122 125 268 308

7

Please reset
combine list before
creating a new list

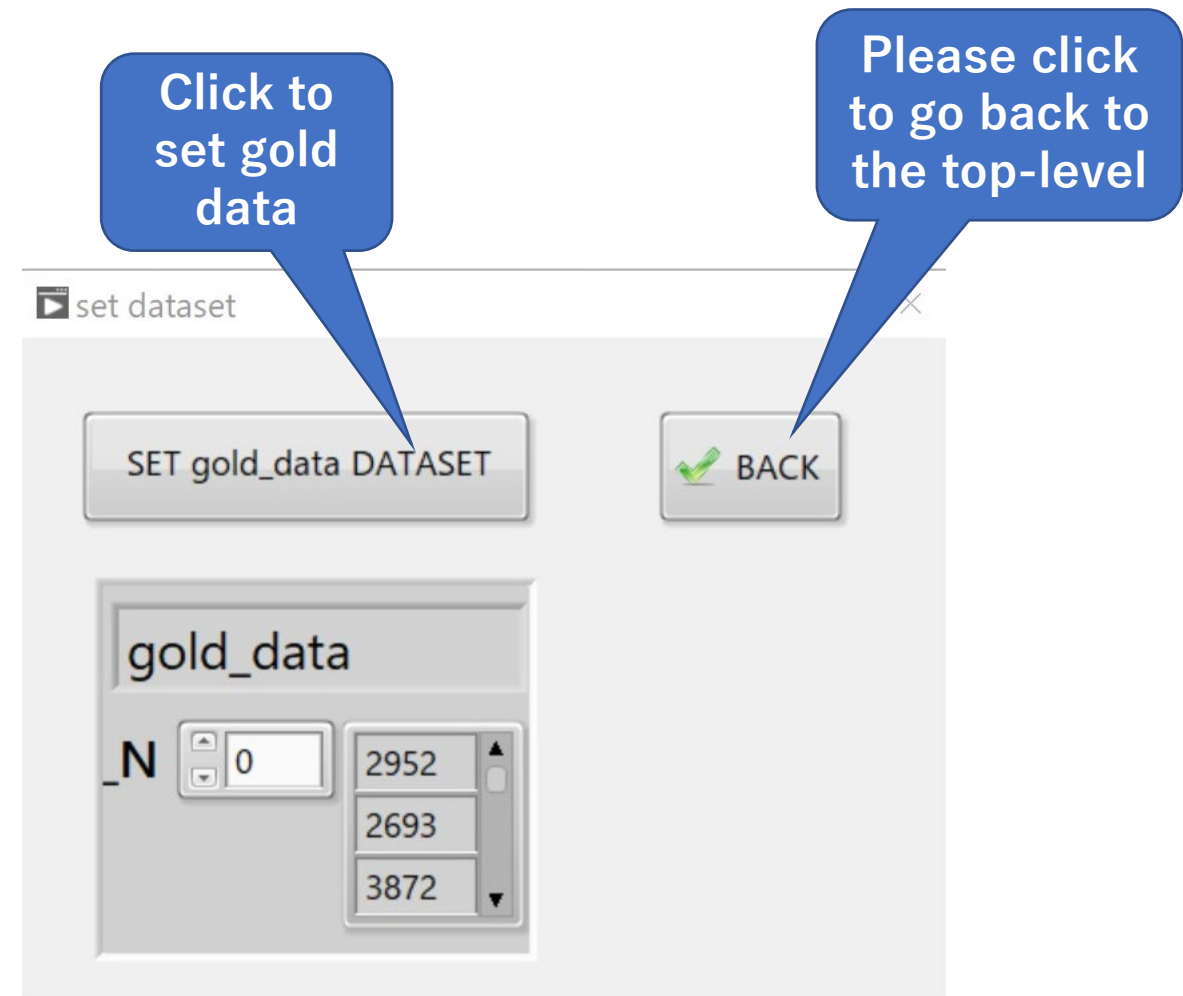
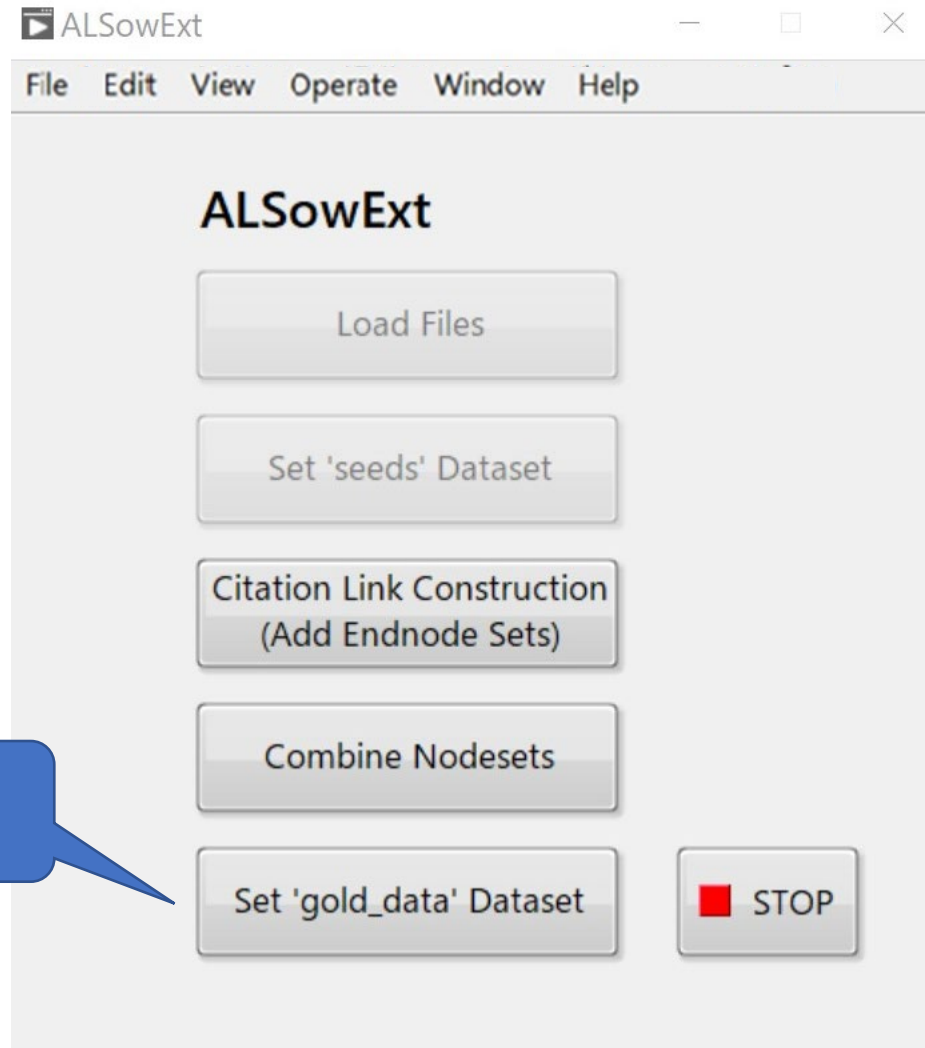
6

Click, then new end-node
sets will be available in
the upper lists

8

Please click
to go back to
the top-level

Setting gold data for the further evaluation



A sample of evaluation with gold data

combine nodeset

Node sets/Endnode of Citation Link

[SF \cup SR] ☒

[SF \cap SR] ☒

gold_data ☒

AND (\cap)

Node sets/Endnode of Citation Link

SR ☒

[SF \cup SR] ☒

[SF \cap SR] ☒

COMBINE

RESET Combine LIST

Count Endnode

Add to Node sets

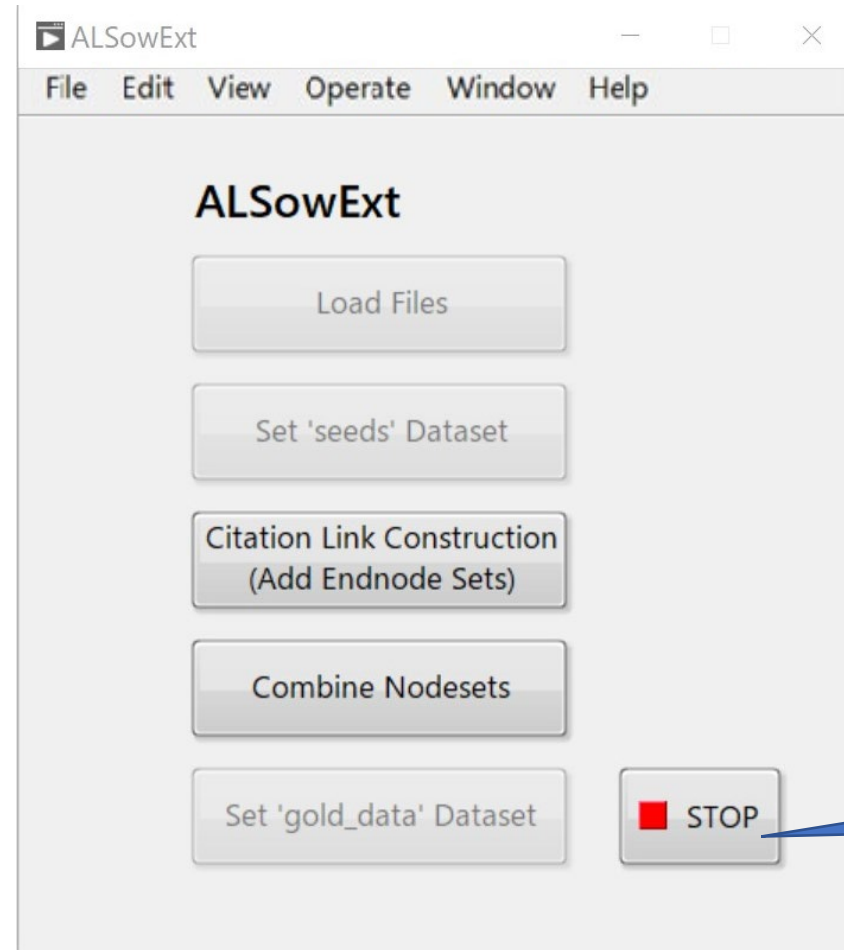
BACK

Combine List

[gold_data \cap [SF \cap SR]]

Combined node set(s)	No. of endnode	ID(_N)
[gold_data \cap [SF \cap SR]]	57	73 125 268 342 471
	0	0 0 0 0 0

Exit of ALSowExt



2

Click to
close

1

Click to stop
the application