▼ 1 Investigating Hybrid Strategies for Systematic Literature Review

Experiment 2

Study

A systematic review on the use of Definition of Done on agile software development projects

Silva, A., Araújo, T., Nunes, J., Perkusich, M.r, Dilorenzo, E., Almeida, H.r, Perkusich, A.
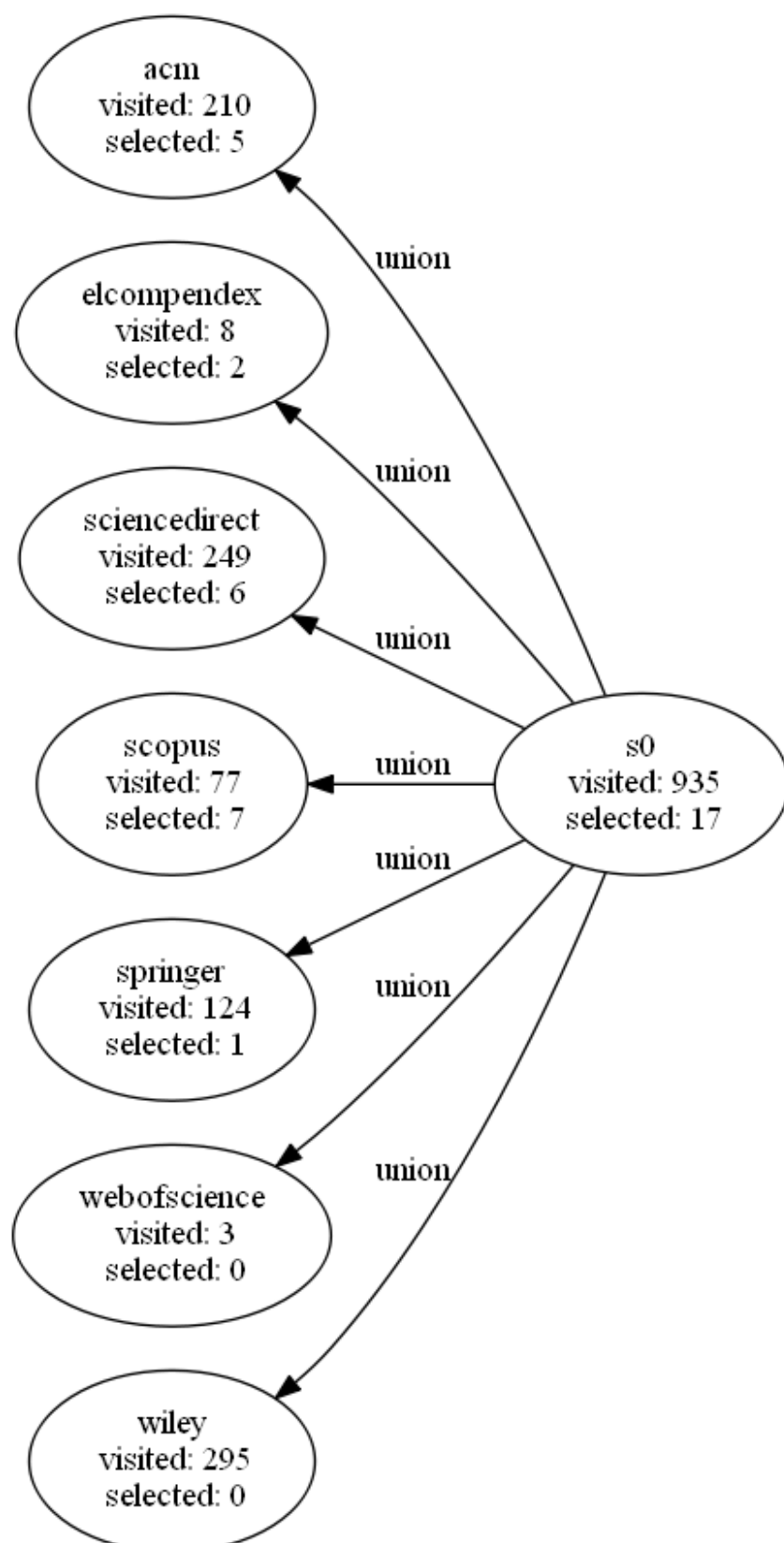
21st International Conference on Evaluation and Assessment in Software Engineering, EASE 2017

▼ 2 Strategies

▼ 2.1 Strategy 1 - Database Search (Search)

```
Total of selected studies = 16 (Seed Set) + 4 (Snowballing) = 20
```

Provenance Graph - Strategy 1 - Database Search (Search)

acm
visited: 210
selected: 5

union

elcompendex
visited: 8
selected: 2

union

sciencedirect
visited: 249
selected: 6

union

scopus
visited: 77
selected: 7

union

s0
visited: 935
selected: 17

union

springer
visited: 124
selected: 1

union

webofscience
visited: 3
selected: 0

union

wiley
visited: 295
selected: 0

Report Strategy 1 - Database Search (Search)

Total of selected studies in the SLR = 20
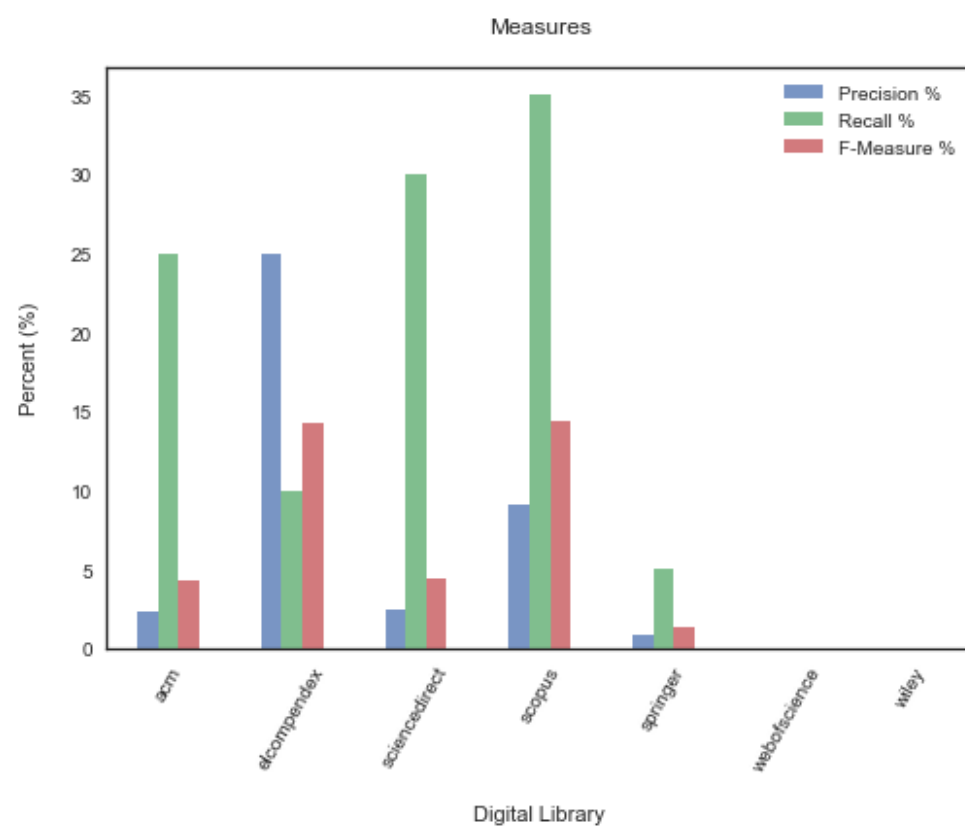
Precision = Selected/ Visited
Recall = Selected/ Total of selected studies in the SLR
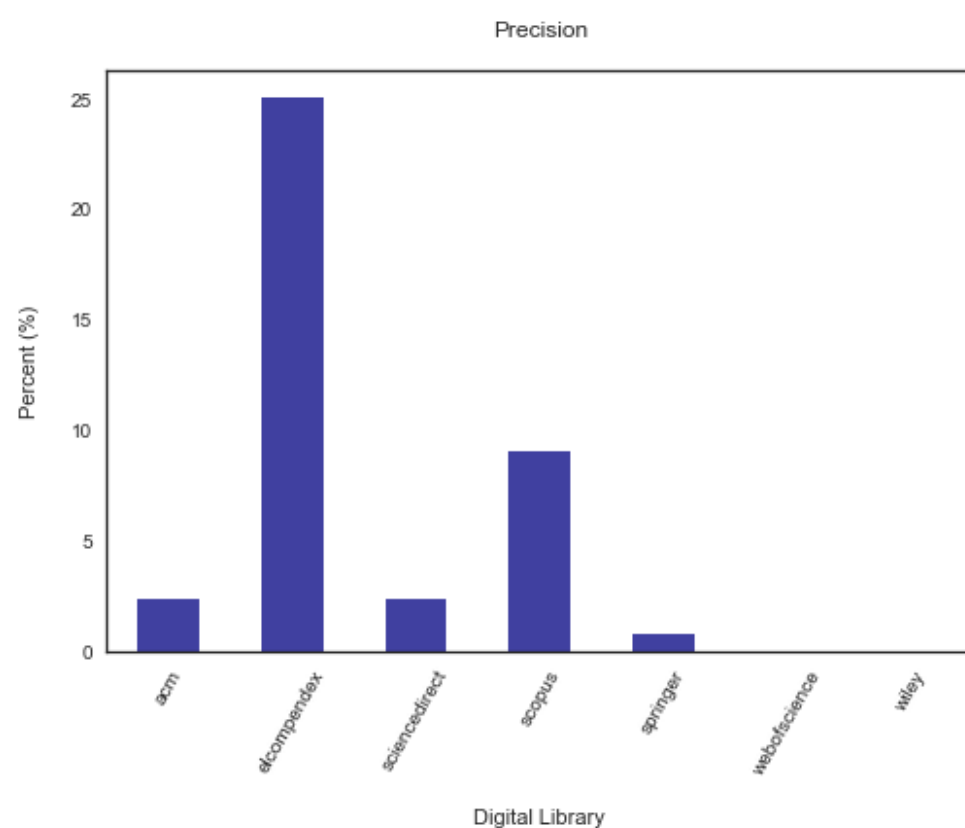F-Measure = Harmonic Mean of Precision and Recall

Out[10]:

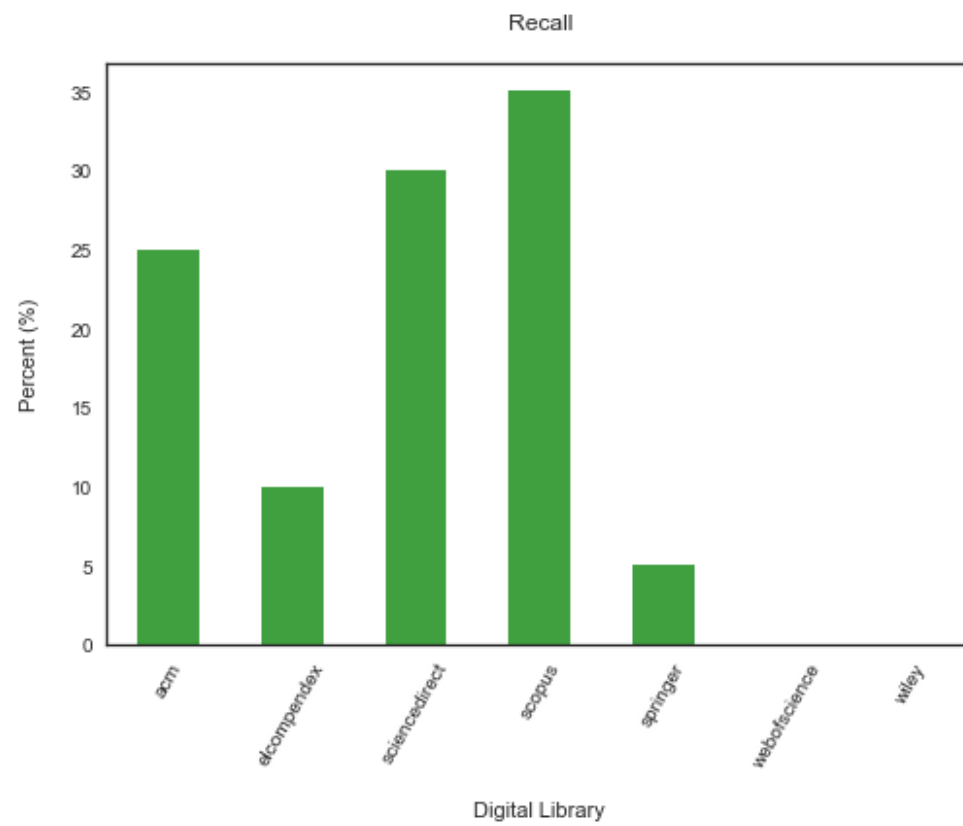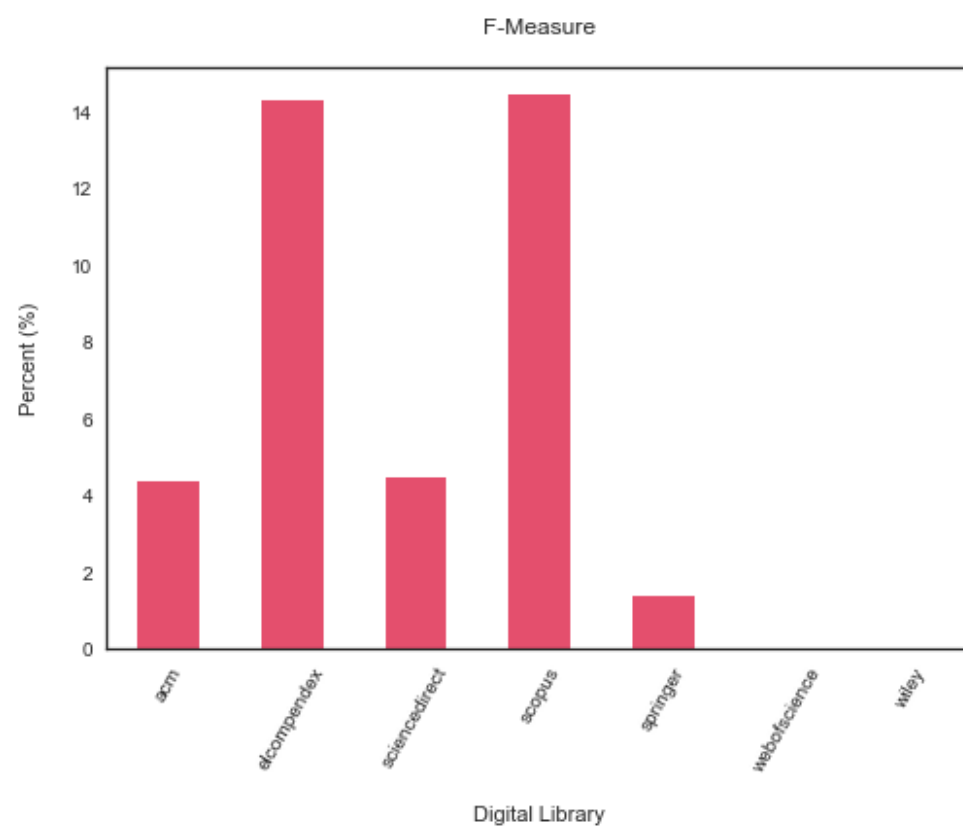| Digital Library	Visited	Selected	Precision %	Recall %	F-Measure %
acm	210	5	2.38	25.00	4.35
elcompendex	8	2	25.00	10.00	14.29
sciencedirect	249	6	2.41	30.00	4.46
scopus	77	7	9.09	35.00	14.43
springer	124	1	0.81	5.00	1.39
webofscience	3	0	0.00	0.00	0.00
wiley	295	0	0.00	0.00	0.00

Bar Chart Strategy 1

Bar Chart Strategy 1

Bar Chart Strategy 1

Recall

Bar Chart Strategy 1

F-Measure

Report Strategy 1 - Measures of Strategy - Union of all digital libraries (accumulated)

Out[16]:

Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Search	935	17	1.82	85.00	3.56

Report Strategy 1 - Total of papers retrieved from Excel and Duplicates.

```
Visited = Excel + Duplicates
Selected = Papers selected in the SLR
Excel =  Papers retrieved from Excel provided by authors
Excel Selected =  Papers selected in the SLR and retrieved from Excel provided by authors
Duplicated =  Papers identified as duplicate
```
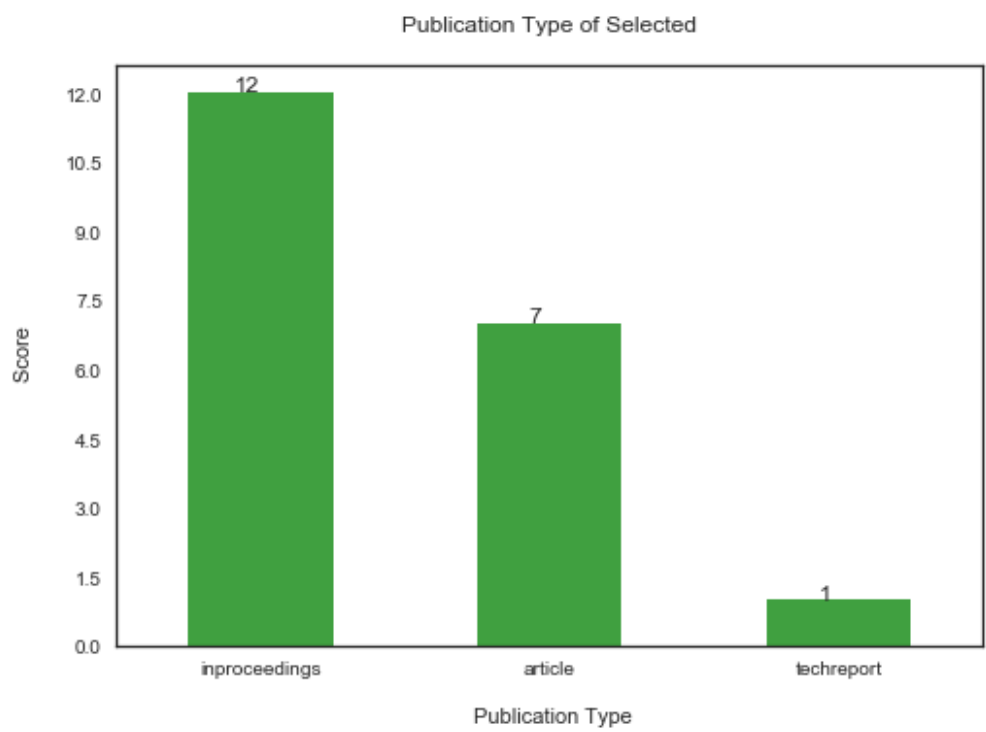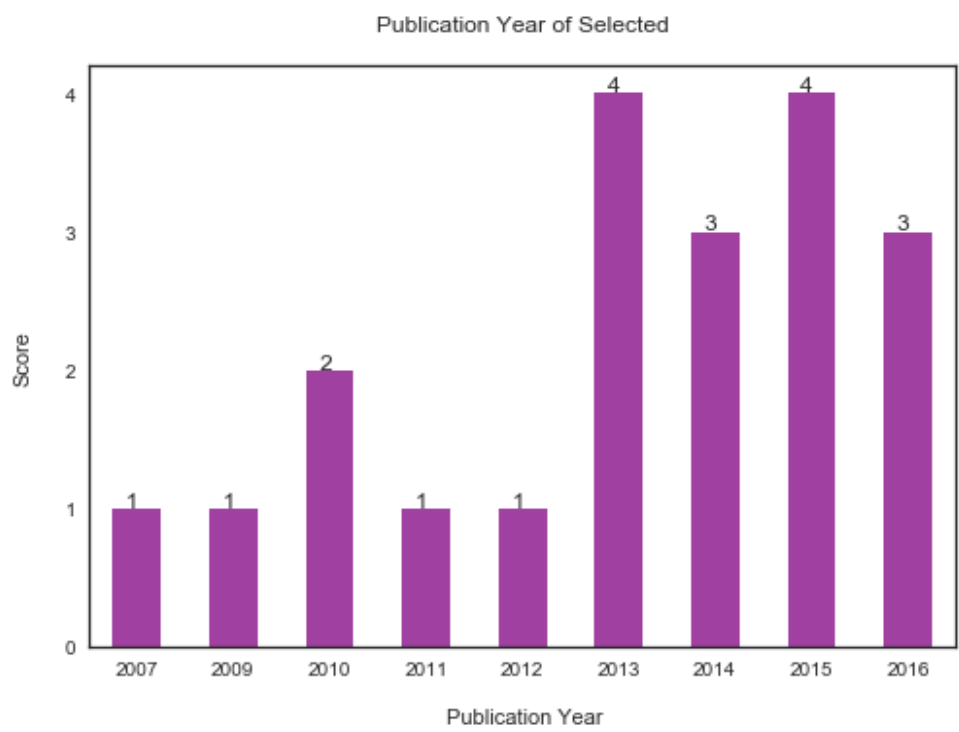
Out[17]:

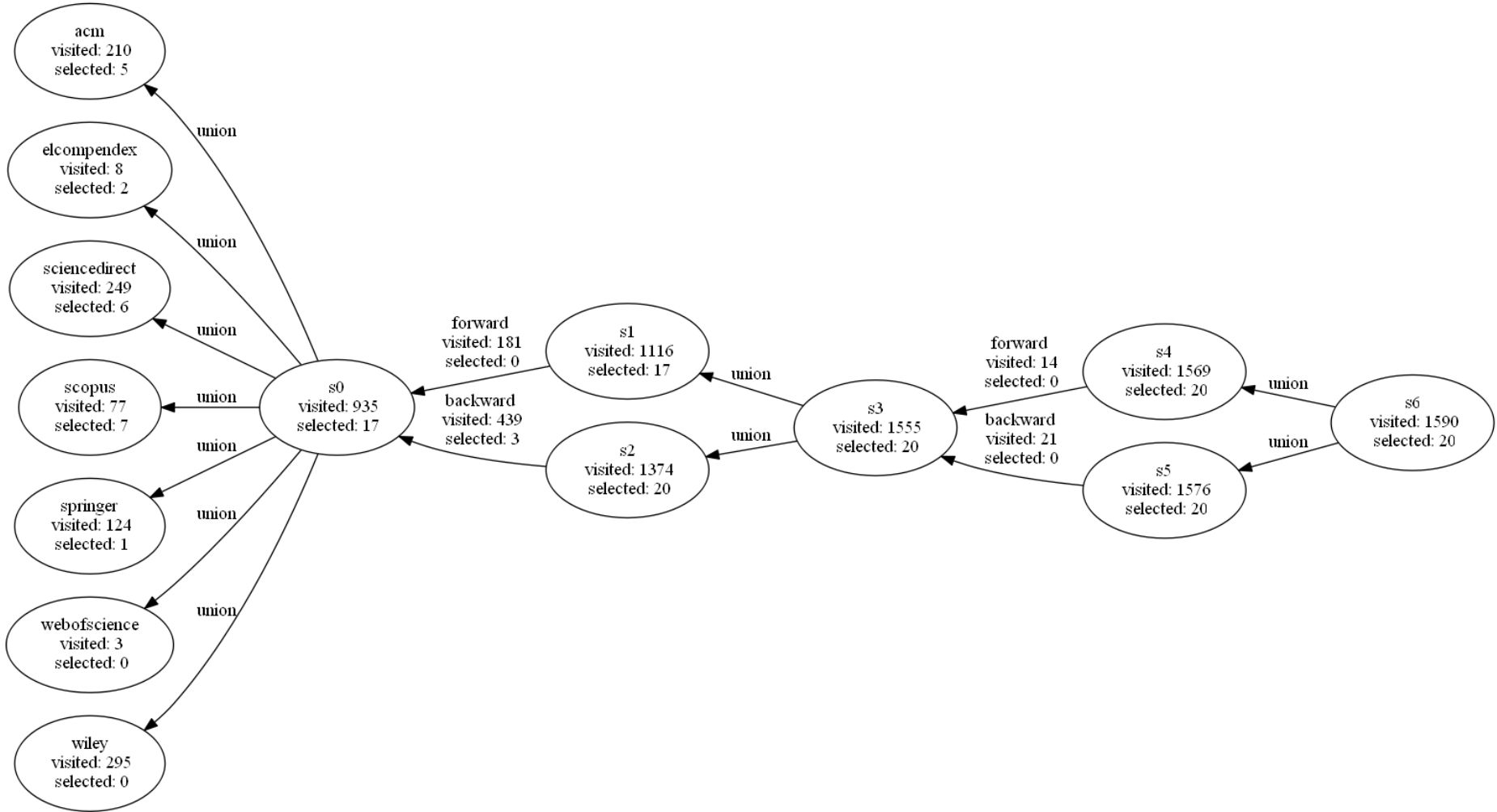| Digital Library	Visited	Selected	Excel	Excel Selected	Duplicated
acm	210	5	210	5	0
elcompendex	8	2	1	0	7
sciencedirect	249	6	245	5	4
scopus	77	7	68	7	9
springer	124	1	115	0	9
webofscience	3	0	3	0	0
wiley	295	0	294	0	1

Bar Chart Publications Type - Strategy 1

Publication Type of Selected

Bar Chart Publications Year - Strategy 1

Publication Year of Selected

▼ 2.2 Strategy 2 - Database Search + Snowballing of Wohlin Guideline (Search + SB)

Provenance Graph - Strategy 2 - Database Search + Snowballing of Wohlin Guideline (Search + SB)

acm
visited: 210
selected: 5

elcompendex
visited: 8
selected: 2

sciencedirect
visited: 249
selected: 6

scopus
visited: 77
selected: 7

springer
visited: 124
selected: 1

webofscience
visited: 3
selected: 0

wiley
visited: 295
selected: 0

s0
visited: 935
selected: 17

union

forward
visited: 181
selected: 0

backward
visited: 439
selected: 3

s1
visited: 1116
selected: 17

s2
visited: 1374
selected: 20

union

s3
visited: 1555
selected: 20

forward
visited: 14
selected: 0

backward
visited: 21
selected: 0

s4
visited: 1569
selected: 20

s5
visited: 1576
selected: 20

union

s6
visited: 1590
selected: 20

Report Strategy 2 - Database Search + Snowballing of Wohlin Guideline (Search + SB)

Total of selected studies in the SLR = 20

Precision = Selected/ Visited
Recall = Selected/ Total of selected studies in the SLR
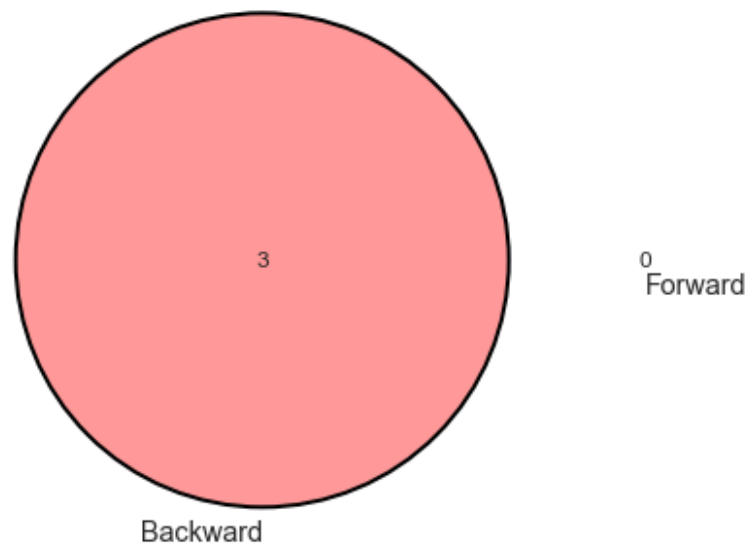F-Measure = Harmonic Mean of Precision and Recall

Out[22]:

Digital Library/State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
acm	210	5	2.38	25.00	4.35	-	210	5	2.38	25.00	4.35
elcompendex	8	2	25.00	10.00	14.29	-	8	2	25.00	10.00	14.29
sciencedirect	249	6	2.41	30.00	4.46	-	249	6	2.41	30.00	4.46
scopus	77	7	9.09	35.00	14.43	-	77	7	9.09	35.00	14.43
springer	124	1	0.81	5.00	1.39	-	124	1	0.81	5.00	1.39
webofscience	3	0	0.00	0.00	0.00	-	3	0	0.00	0.00	0.00
wiley	295	0	0.00	0.00	0.00	-	295	0	0.00	0.00	0.00
s0	0	0	nan	0.00	0.00	union	935	17	1.82	85.00	3.56
s1	181	0	0.00	0.00	0.00	forward	1116	17	1.52	85.00	2.99
s2	439	3	0.68	15.00	1.31	backward	1374	20	1.46	100.00	2.87
s3	0	0	nan	0.00	0.00	union	1555	20	1.29	100.00	2.54
s4	14	0	0.00	0.00	0.00	forward	1569	20	1.27	100.00	2.52
s5	21	0	0.00	0.00	0.00	backward	1576	20	1.27	100.00	2.51
s6	0	0	nan	0.00	0.00	union	1590	20	1.26	100.00	2.48

Report Strategy 2 - Measures of Strategy (accumulated)

Out[23]:

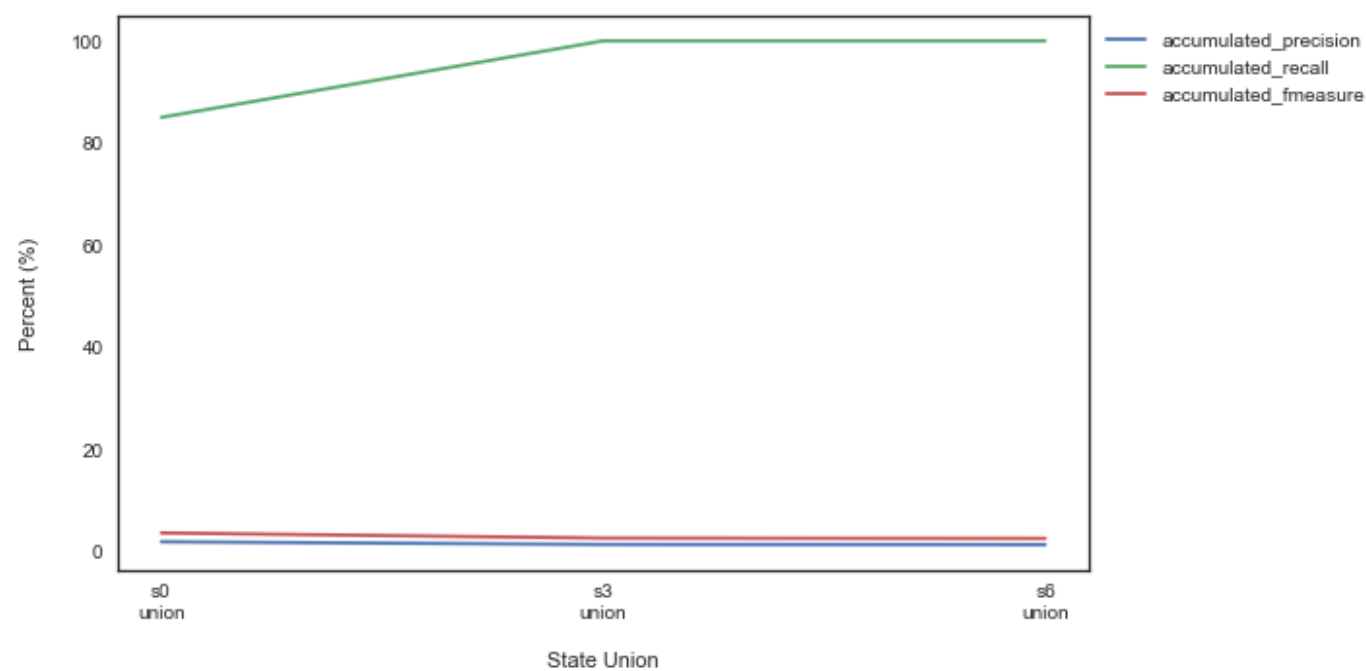| Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Search + SB	1590	20	1.26	100.00	2.48

Diagrama de Venn - Strategy 2
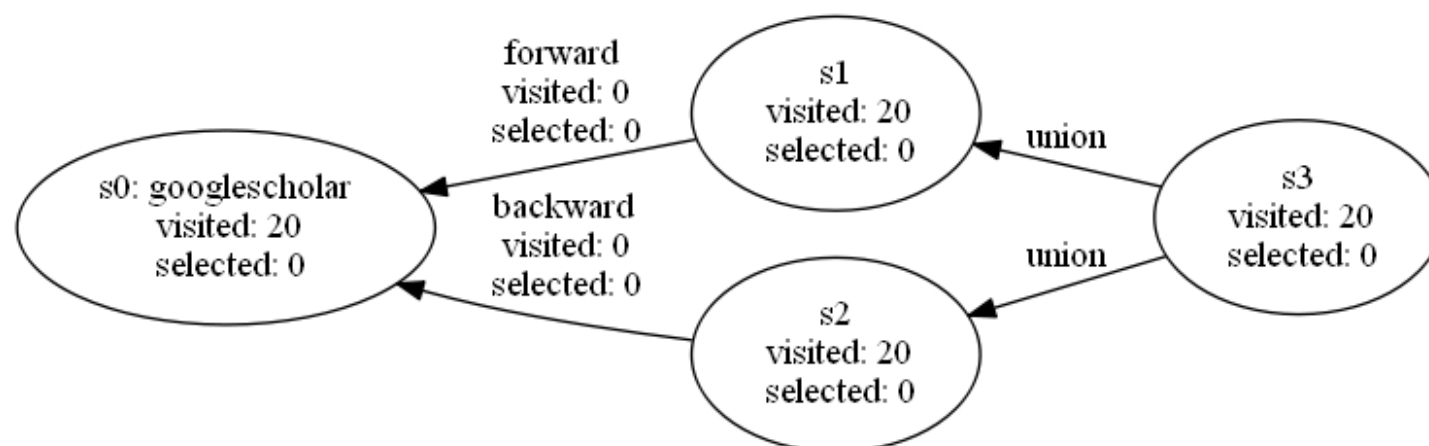
Report Diagrama de Venn

Backward ['barton2009a', 'saddington2012a', 'eloranta2013a']
Forward []

Accumulated Measures - Strategy 2

▼ 2.3 Strategy 3 - Informal Search in Google Scholar + Snowballing of Wohlin Guideline = (GS + SB)

Provenance Graph - Strategy 3 - Informal Search in Google Scholar + Snowballing of Wohlin Guideline = (GS + SB)

Report Strategy 3 - Informal Search in Google Scholar + Snowballing of Wohlin Guideline = (GS + SB)

Total of selected studies in the SLR = 20

Precision = Selected/ Visited
Recall = Selected/ Total of selected studies in the SLR
F-Measure = Harmonic Mean of Precision and Recall

Out[29]:

Scholar=s0 /State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
s0: googlescholar	20	0	0.00	0.00	0.00	-	20	0	0.00	0.00	0.00
s1	0	0	nan	0.00	0.00	forward	20	0	0.00	0.00	0.00
s2	0	0	nan	0.00	0.00	backward	20	0	0.00	0.00	0.00
s3	0	0	nan	0.00	0.00	union	20	0	0.00	0.00	0.00

Report Strategy 3 - Measures of Strategy (accumulated)

Out[30]:

Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
GS + SB	20	0	0.00	0.00	0.00

C:\ProgramData\Anaconda3\lib\site-packages\matplotlib_venn_venn2.py:50: UserWarning: Both circles have zero area
  warnings.warn("Both circles have zero area")
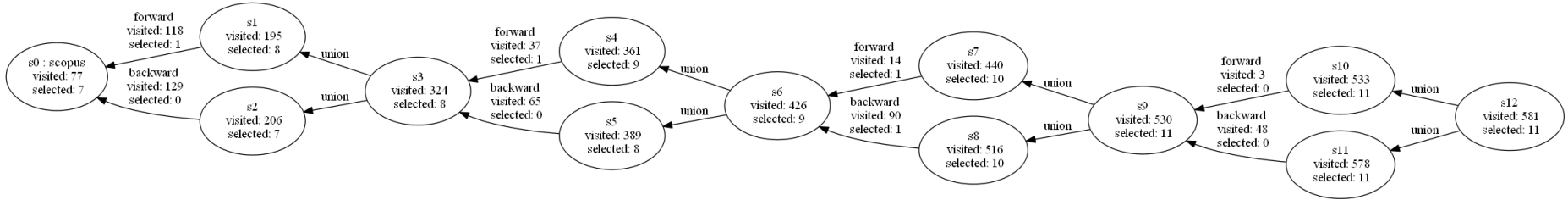
Diagrama de Venn - Strategy 3

Backward[0]   [0]Forward

Report Diagrama de Venn

Backward []
Forward []

2.4 Strategy 4 - Scopus + Snowballing of Wohlin Guideline (Scopus + SB)

Provenance Graph - Strategy 4 - Scopus + Snowballing of Wohlin Guideline (Scopus + SB)

Report Strategy 4 - Scopus + Snowballing of Wohlin Guideline (Scopus + SB)

```
Total of selected studies in the SLR =  20

Precision = Selected/ Visited
Recall = Selected/ Total of selected studies in the SLR
F-Measure = Harmonic Mean of Precision and Recall
```
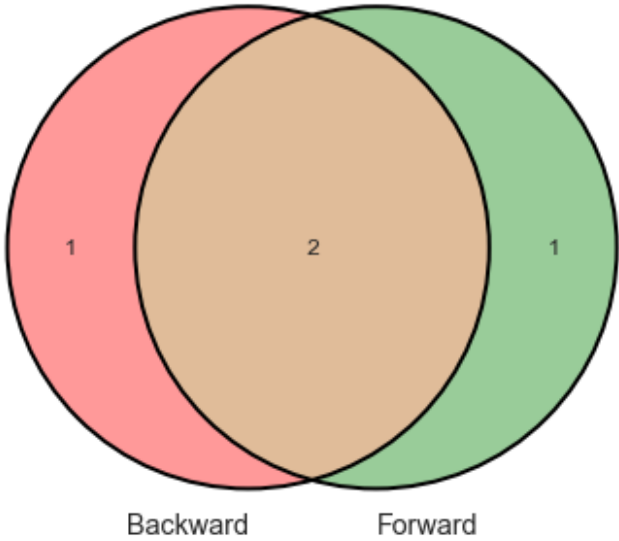
Out[35]:

Scopus=s0 /State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
s0 : scopus	77	7	9.09	35.00	14.43	-	77	7	9.09	35.00	14.43
s1	118	1	0.85	5.00	1.45	forward	195	8	4.10	40.00	7.44
s2	129	0	0.00	0.00	0.00	backward	206	7	3.40	35.00	6.19
s3	0	0	nan	0.00	0.00	union	324	8	2.47	40.00	4.65
s4	37	1	2.70	5.00	3.51	forward	361	9	2.49	45.00	4.72
s5	65	0	0.00	0.00	0.00	backward	389	8	2.06	40.00	3.91
s6	0	0	nan	0.00	0.00	union	426	9	2.11	45.00	4.04
s7	14	1	7.14	5.00	5.88	forward	440	10	2.27	50.00	4.35
s8	90	1	1.11	5.00	1.82	backward	516	10	1.94	50.00	3.73
s9	0	0	nan	0.00	0.00	union	530	11	2.08	55.00	4.00
s10	3	0	0.00	0.00	0.00	forward	533	11	2.06	55.00	3.98
s11	48	0	0.00	0.00	0.00	backward	578	11	1.90	55.00	3.68
s12	0	0	nan	0.00	0.00	union	581	11	1.89	55.00	3.66

Report Strategy 4 - Measures of Strategy (accumulated)

Out[36]:

Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Scopus + SB	581	11	1.89	55.00	3.66

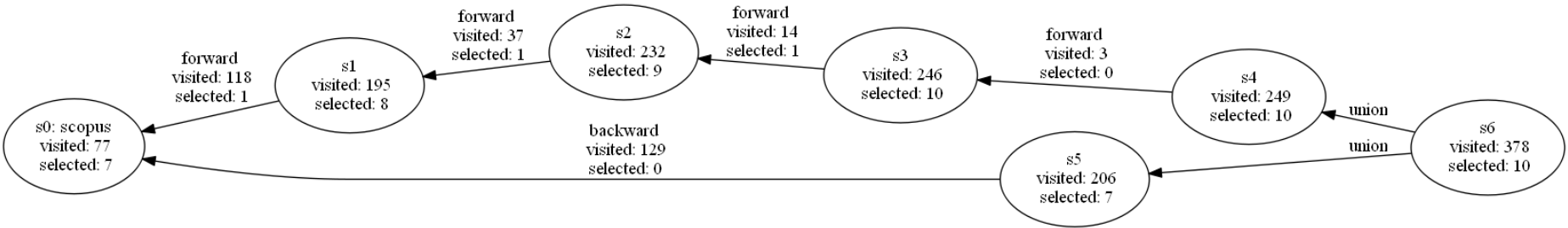Diagrama de Venn - Strategy 4

Report Diagrama de Venn

```
Backward ['waardenburg2013a', 'saddington2012a', 'vlietland2015a']
Forward ['waardenburg2013a', 'vlietland2016a', 'vlietland2015a']
```

▼ 2.5 Strategy 5 - Scopus + Parallel Snowballing of Paper ESEM2017 (Scopus + BS // FS)

Provenance Graph - Strategy 5 - Scopus + Parallel Snowballing of Paper ESEM2017 (Scopus + BS // FS)

Report Strategy 5 - Scopus + Parallel Snowballing of Paper ESEM2017 (Scopus + BS // FS)

Total of selected studies in the SLR = 20

Precision = Selected/ Visited
Recall = Selected/ Total of selected studies in the SLR
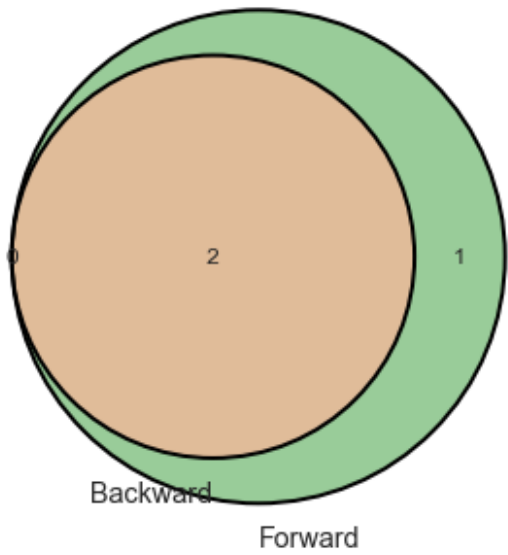F-Measure = Harmonic Mean of Precision and Recall

Out[41]:

Scopus=s0 /State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
s0: scopus	77	7	9.09	35.00	14.43	-	77	7	9.09	35.00	14.43
s1	118	1	0.85	5.00	1.45	forward	195	8	4.10	40.00	7.44
s2	37	1	2.70	5.00	3.51	forward	232	9	3.88	45.00	7.14
s3	14	1	7.14	5.00	5.88	forward	246	10	4.07	50.00	7.52
s4	3	0	0.00	0.00	0.00	forward	249	10	4.02	50.00	7.43
s5	129	0	0.00	0.00	0.00	backward	206	7	3.40	35.00	6.19
s6	0	0	nan	0.00	0.00	union	378	10	2.65	50.00	5.03

Report Strategy 5 - Measures of Strategy (accumulated)

Out[42]:

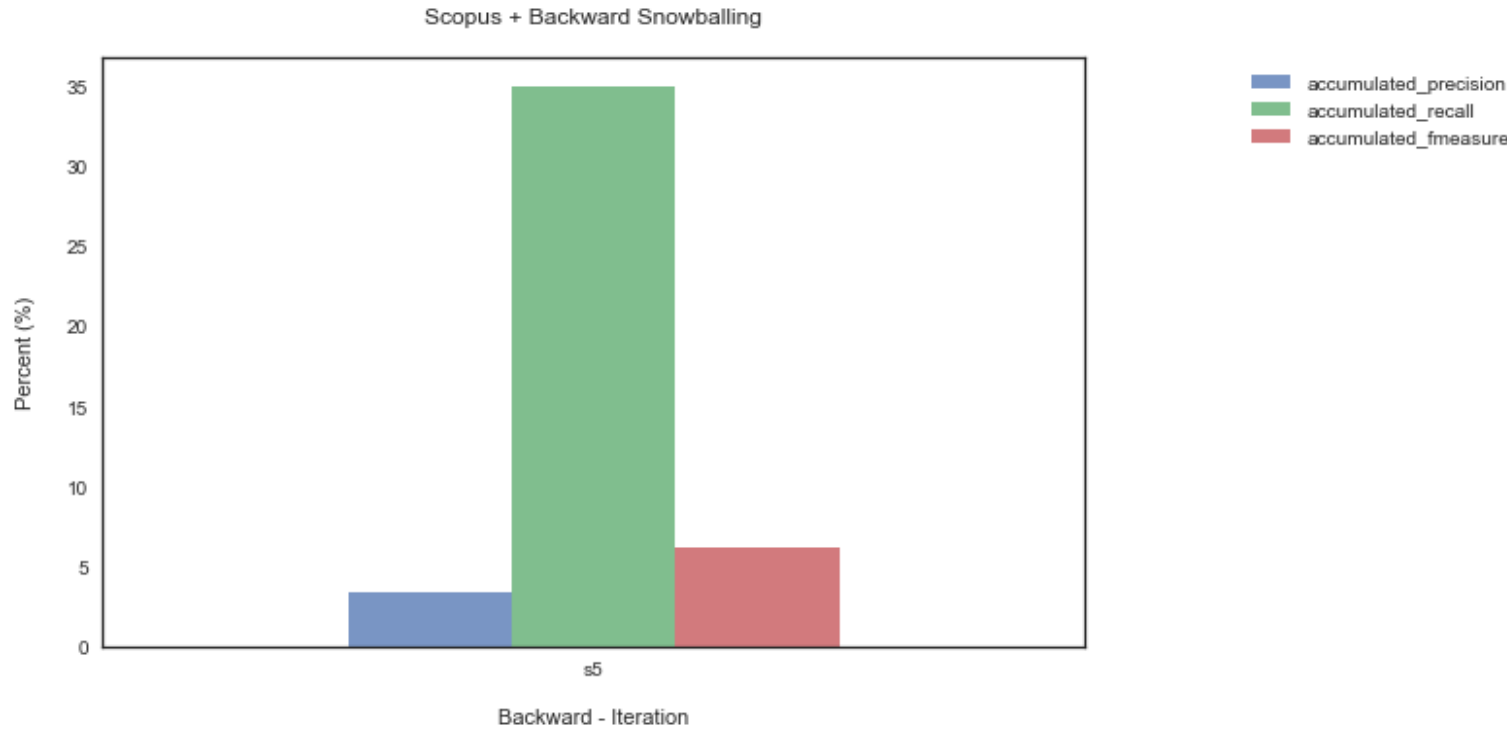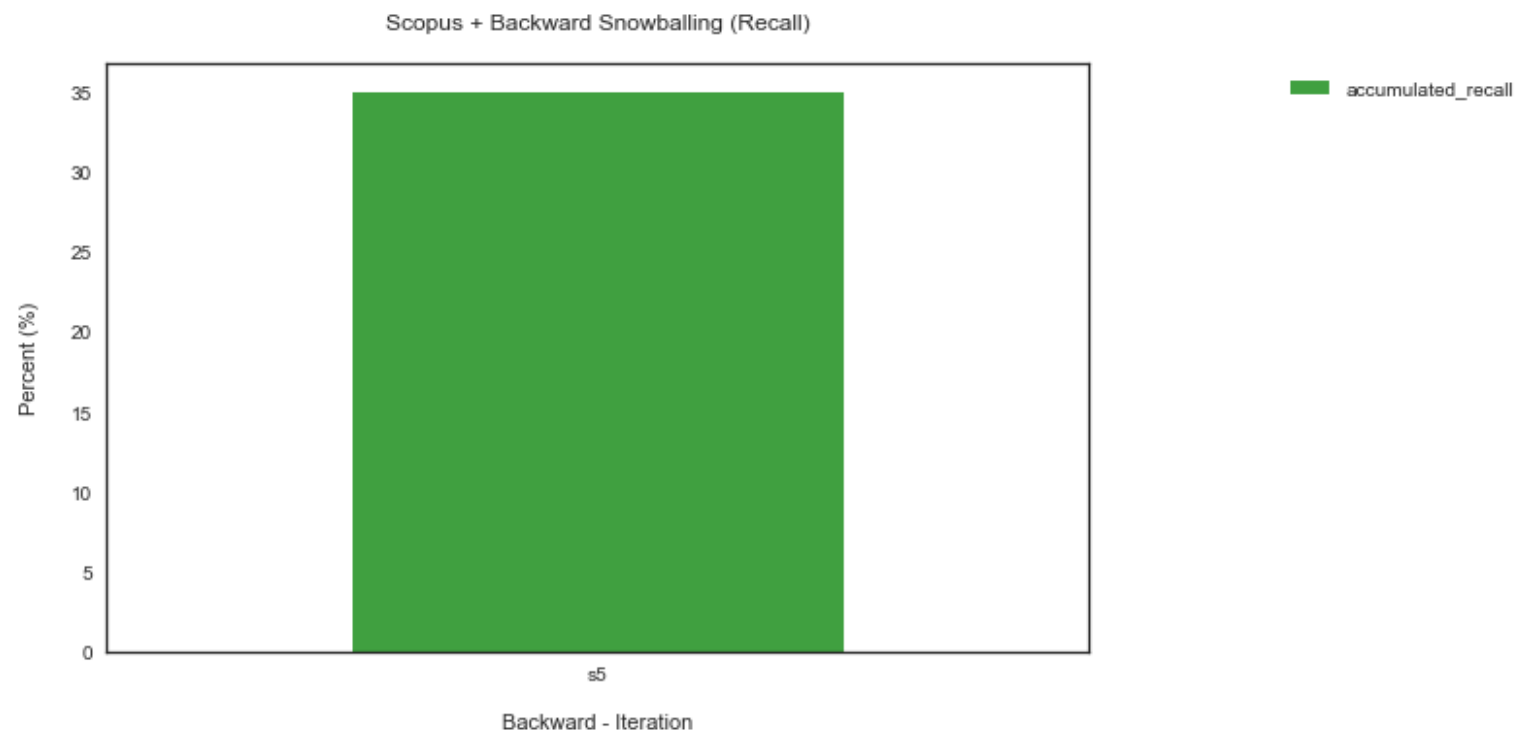| Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Scopus + BS // FS	378	10	2.65	50.00	5.03

Diagrama de Venn - Strategy 5

Report Diagrama de Venn

Backward ['waardenburg2013a', 'vlietland2015a']
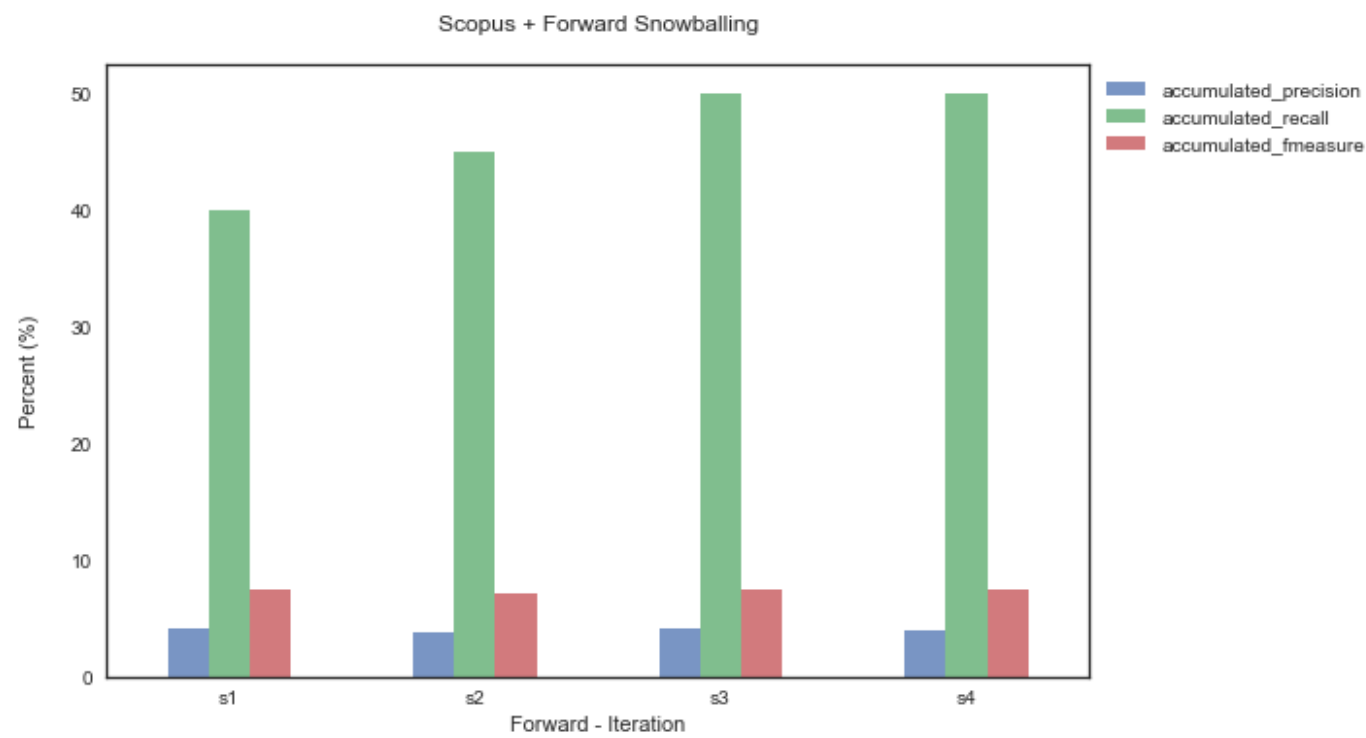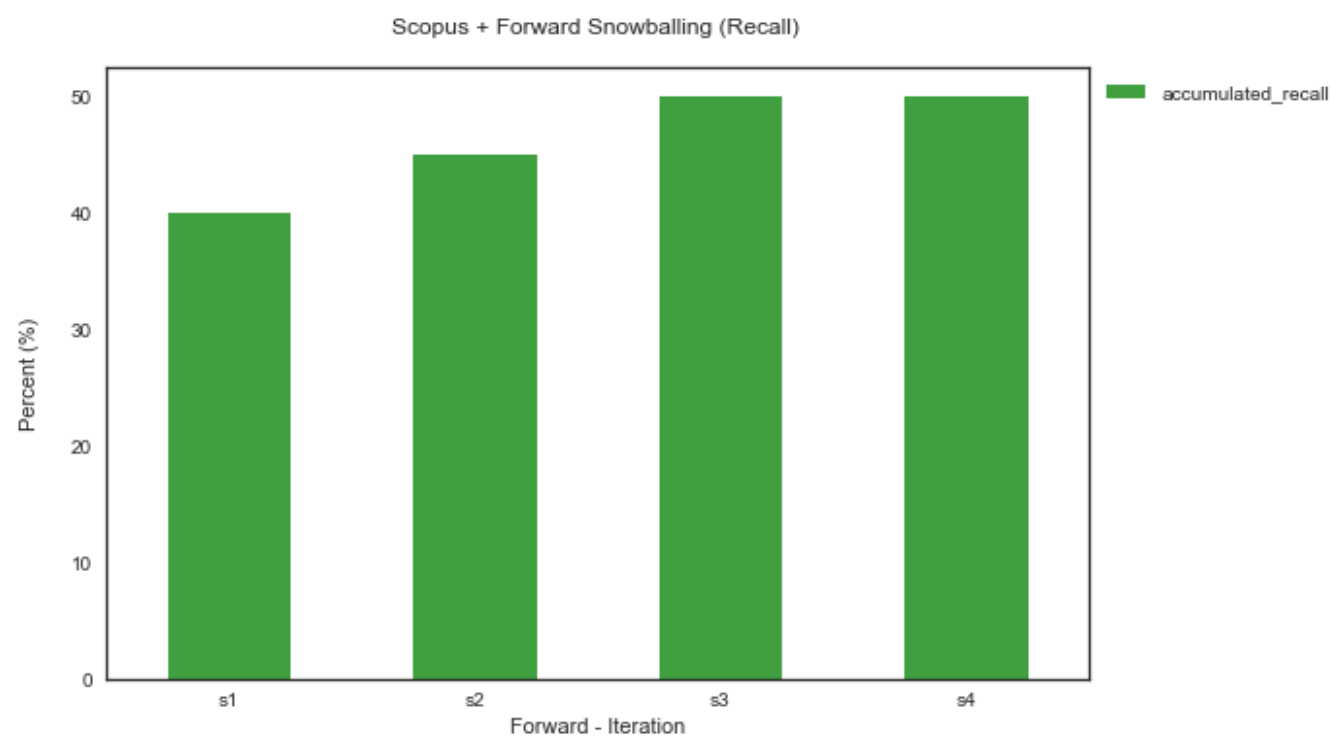Forward ['waardenburg2013a', 'vlietland2016a', 'vlietland2015a']
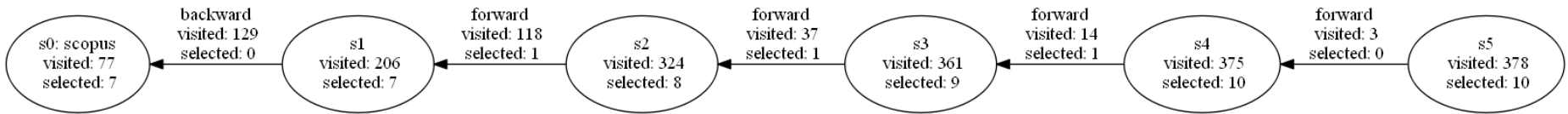
Bar Chart Strategy 5

Scopus + Backward Snowballing

Bar Chart Strategy 5

Scopus + Backward Snowballing (Recall)

Bar Chart Strategy 5

Scopus + Forward Snowballing

Bar Chart Strategy 5

Scopus + Forward Snowballing (Recall)

▼ **2.6 Strategy 6 - Scopus + Sequential Snowballing (Scopus + BS + FS)**

Report Strategy 6 - Scopus + Sequential Snowballing (Scopus + BS + FS)

Total of selected studies in the SLR = 20

Precision = Selected/ Visited
Recall = Selected/ Total of selected studies in the SLR
F-Measure = Harmonic Mean of Precision and Recall
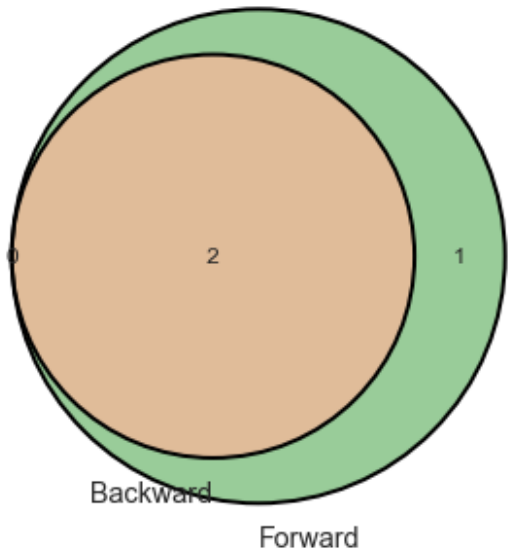
Out[51]:

Scopus=s0 /State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
s0: scopus	77	7	9.09	35.00	14.43	-	77	7	9.09	35.00	14.43
s1	129	0	0.00	0.00	0.00	backward	206	7	3.40	35.00	6.19
s2	118	1	0.85	5.00	1.45	forward	324	8	2.47	40.00	4.65
s3	37	1	2.70	5.00	3.51	forward	361	9	2.49	45.00	4.72
s4	14	1	7.14	5.00	5.88	forward	375	10	2.67	50.00	5.06
s5	3	0	0.00	0.00	0.00	forward	378	10	2.65	50.00	5.03

Report Strategy 6 - Measures of Strategy (accumulated)

Out[52]:

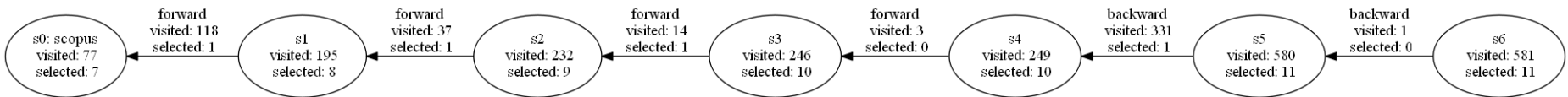| Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Scopus + BS + FS	378	10	2.65	50.00	5.03

Diagrama de Venn - Strategy 6

Report Diagrama de Venn

Backward ['waardenburg2013a', 'vlietland2015a']
Forward ['waardenburg2013a', 'vlietland2016a', 'vlietland2015a']

▼ 2.7 Strategy 7 - Scopus + Sequential Snowballing (Scopus + FS + BS)

Report Strategy 7 - Scopus + Sequential Snowballing (Scopus + FS + BS)

Total of selected studies in the SLR = 20

Precision = Selected/ Visited
Recall = Selected/ Total of selected studies in the SLR
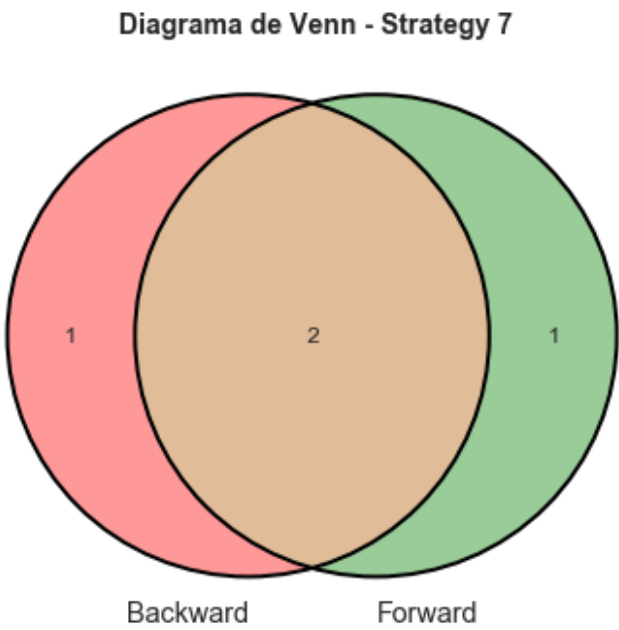F-Measure = Harmonic Mean of Precision and Recall

Out[57]:

Scopus=s0 /State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
s0: scopus	77	7	9.09	35.00	14.43	-	77	7	9.09	35.00	14.43
s1	118	1	0.85	5.00	1.45	forward	195	8	4.10	40.00	7.44
s2	37	1	2.70	5.00	3.51	forward	232	9	3.88	45.00	7.14
s3	14	1	7.14	5.00	5.88	forward	246	10	4.07	50.00	7.52
s4	3	0	0.00	0.00	0.00	forward	249	10	4.02	50.00	7.43
s5	331	1	0.30	5.00	0.57	backward	580	11	1.90	55.00	3.67
s6	1	0	0.00	0.00	0.00	backward	581	11	1.89	55.00	3.66

Report Strategy 7 - Measures of Strategy (accumulated)

Out[58]:

Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Scopus + FS + BS	581	11	1.89	55.00	3.66

Diagrama de Venn - Strategy 7
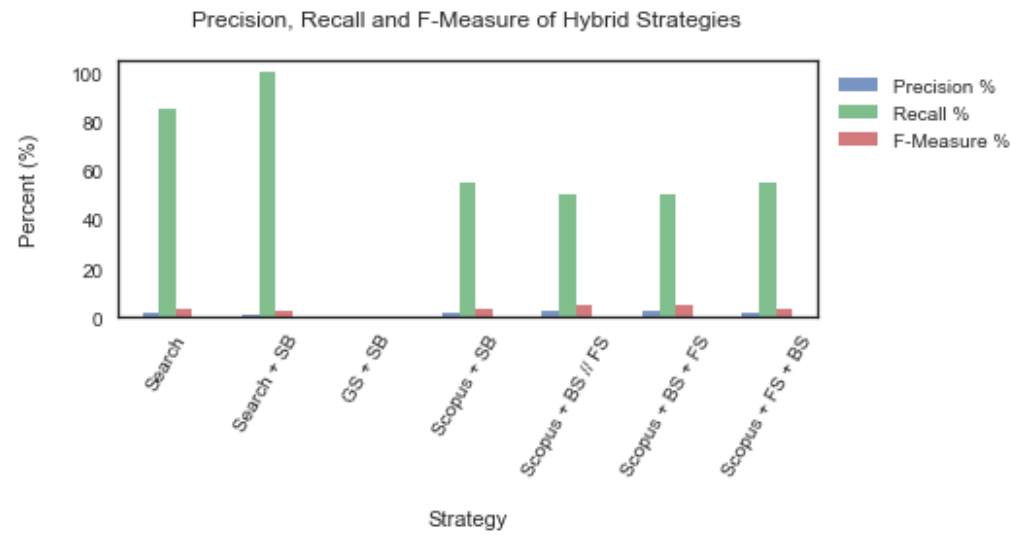
Report Diagrama de Venn

Backward ['waardenburg2013a', 'saddington2012a', 'vlietland2015a']
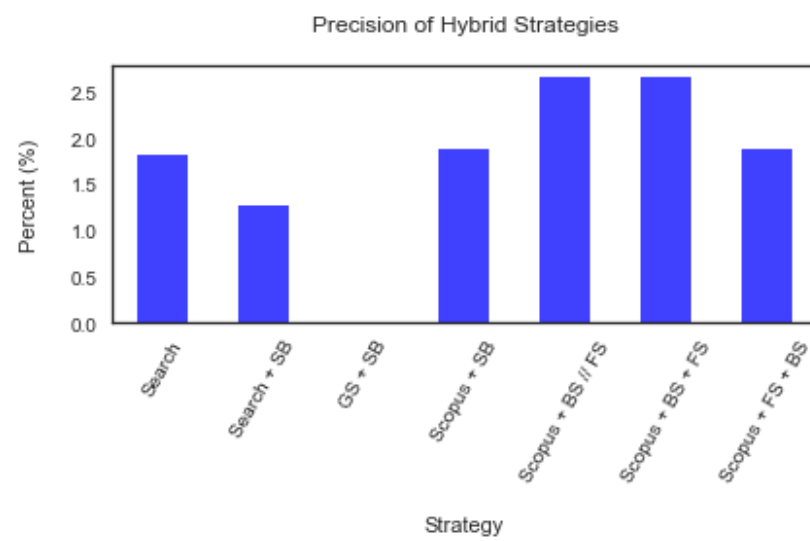Forward ['waardenburg2013a', 'vlietland2016a', 'vlietland2015a']

▼ 3 Analyses

Out[62]:

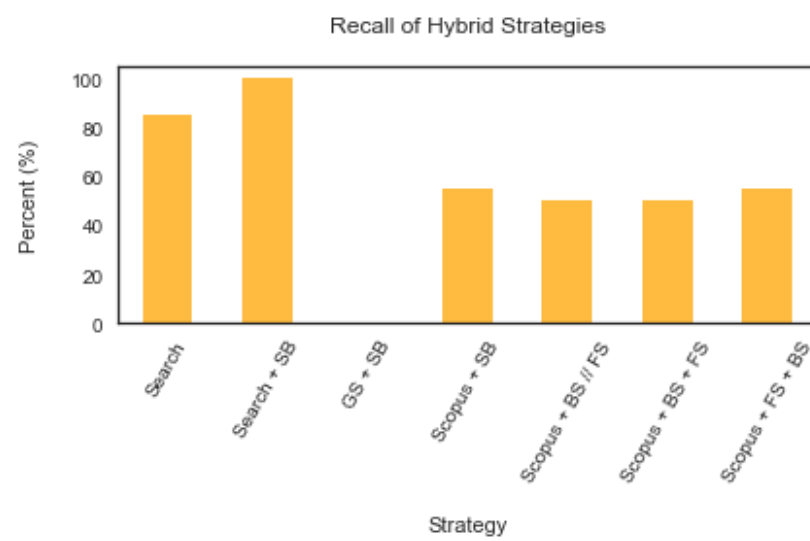| Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Search	935	17	1.82	85.00	3.56
Search + SB	1590	20	1.26	100.00	2.48
GS + SB	20	0	0.00	0.00	0.00
Scopus + SB	581	11	1.89	55.00	3.66
Scopus + BS // FS	378	10	2.65	50.00	5.03
Scopus + BS + FS	378	10	2.65	50.00	5.03
Scopus + FS + BS	581	11	1.89	55.00	3.66

Bar Chart Strategies

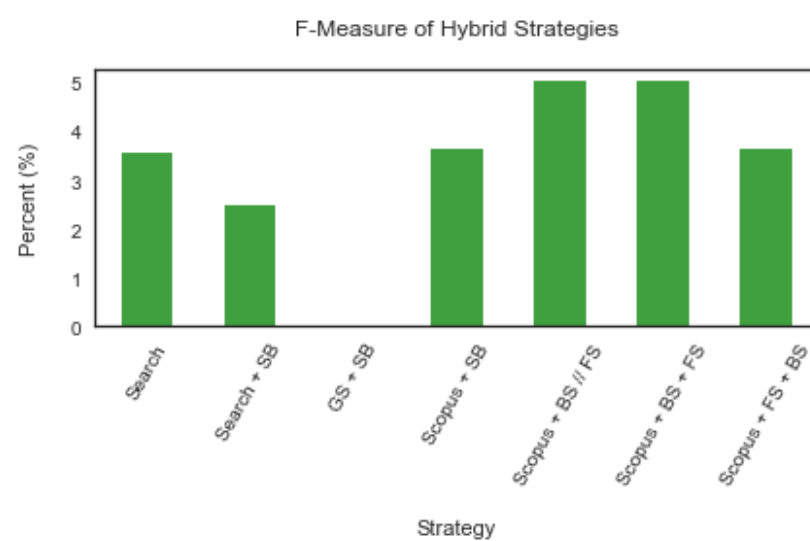Precision, Recall and F-Measure of Hybrid Strategies

Bar Chart Strategies

Precision of Hybrid Strategies

Bar Chart Strategies

Recall of Hybrid Strategies

Bar Chart Strategies

F-Measure of Hybrid Strategies

▼ 4 Research Questions

RQ1) Which of the digital libraries used in the published SLR is more efficient?

Out[69]:

Digital Library	Visited	Selected	Precision %	Recall %	F-Measure %
acm	210	5	2.38	25.00	4.35
elcompendex	8	2	25.00	10.00	14.29
sciencedirect	249	6	2.41	30.00	4.46
scopus	77	7	9.09	35.00	14.43
springer	124	1	0.81	5.00	1.39
webofscience	3	0	0.00	0.00	0.00
wiley	295	0	0.00	0.00	0.00

RQ2) Which of the hybrid strategies is more efficient?

Out[70]:

Strategy	Precision %	Recall %	F-Measure %
Search + SB	1.26	100.00	2.48
GS + SB	0.00	0.00	0.00
Scopus + SB	1.89	55.00	3.66
Scopus + BS // FS	2.65	50.00	5.03
Scopus + BS + FS	2.65	50.00	5.03
Scopus + FS + BS	1.89	55.00	3.66

RQ3) Which of the strategies is more efficient?

Out[71]:

Strategy	Precision %	Recall %	F-Measure %
Search	1.82	85.00	3.56
Search + SB	1.26	100.00	2.48
GS + SB	0.00	0.00	0.00
Scopus + SB	1.89	55.00	3.66
Scopus + BS // FS	2.65	50.00	5.03
Scopus + BS + FS	2.65	50.00	5.03
Scopus + FS + BS	1.89	55.00	3.66

RQ4) How many selected articles does the digital library find through the direct search for the title of the article?

Out[72]:

Digital Library	Recall %
acm	25.00
elcompendex	95.00
sciencedirect	30.00
scopus	95.00
springer	10.00
webofscience	55.00
wiley	0.00