

## ▼ 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, Dilozenzo, E., Almeida, H.r, Perkusich, A.

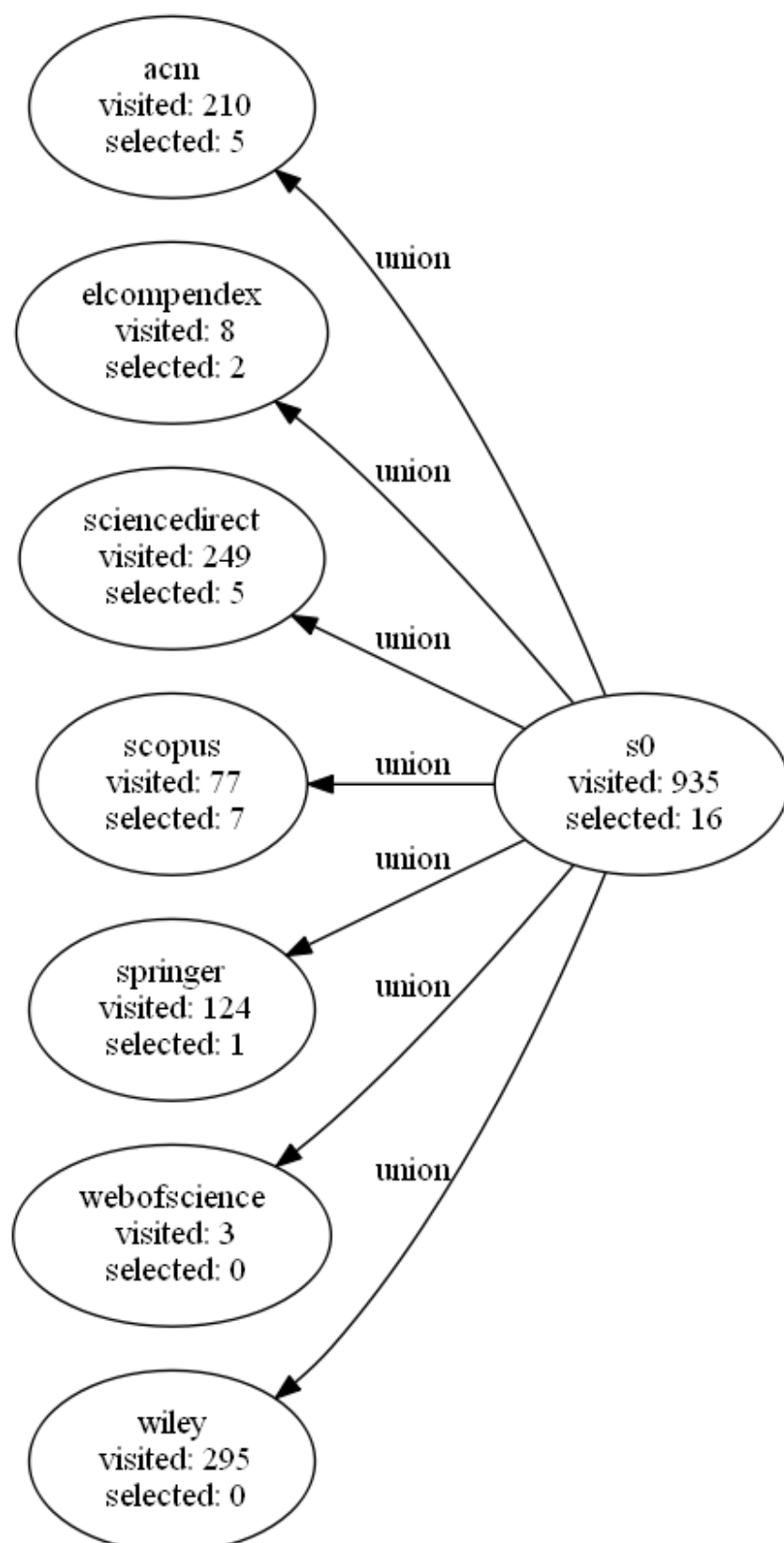
21st International Conference on Evaluation and Assessment in Software Engineering, EASE 2017

## ▼ 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)



## Report Strategy 1 - Database Search (Search)

Total of selected studies in the SLR = 20

Precision = Selected/ Visited

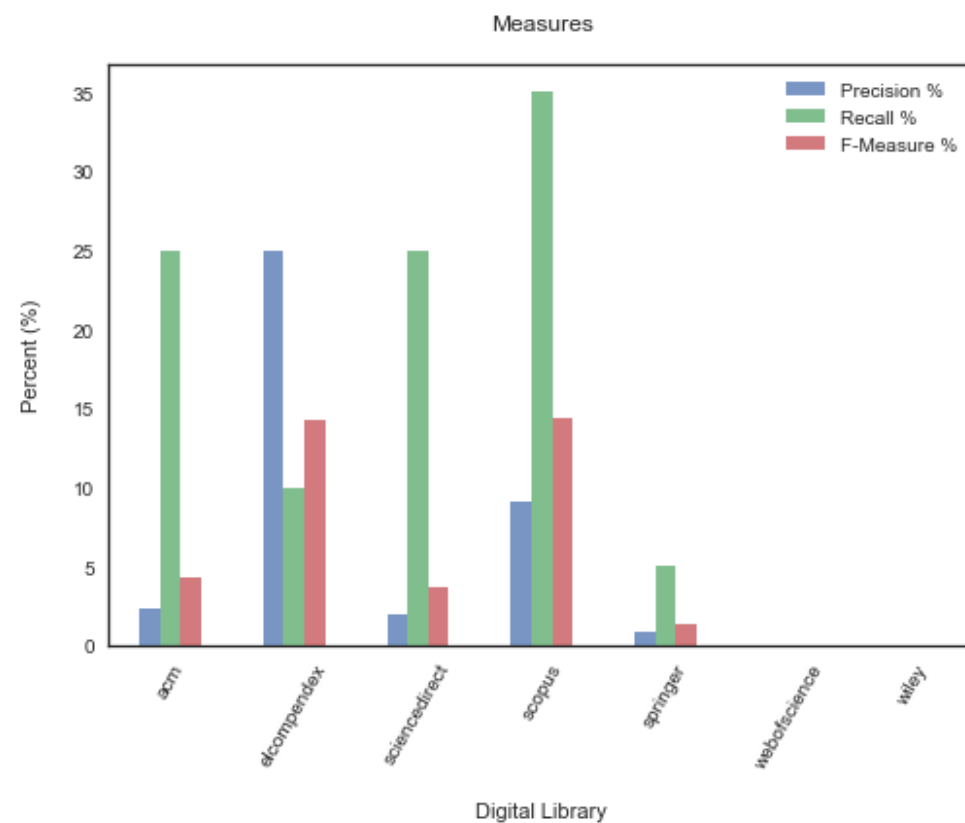
Recall = Selected/ Total of selected studies in the SLR

F-Measure = Harmonic Mean of Precision and Recall

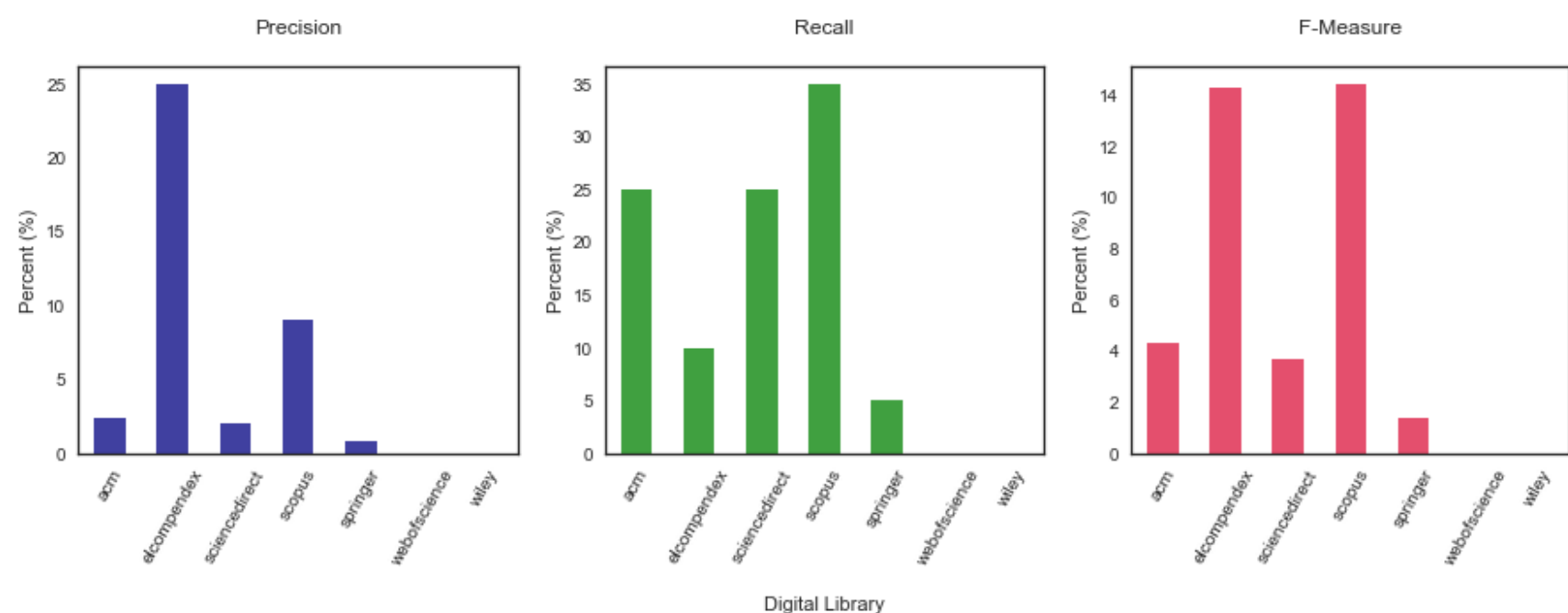
Out[10]:

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	5	2.01	25.00	3.72
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



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

Out[17]:

Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Search	935	16	1.71	80.00	3.35

# 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

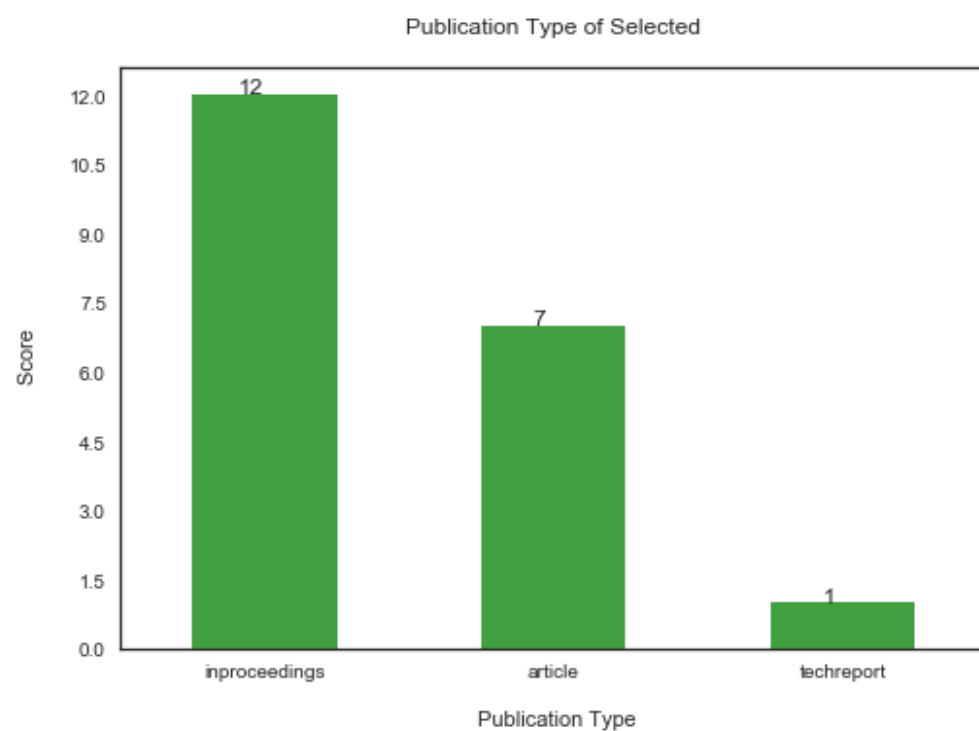
Excel Selected = Papers selected in the SLR and retrieved from Excel provided by authors

Duplicated = Papers identified as duplicate

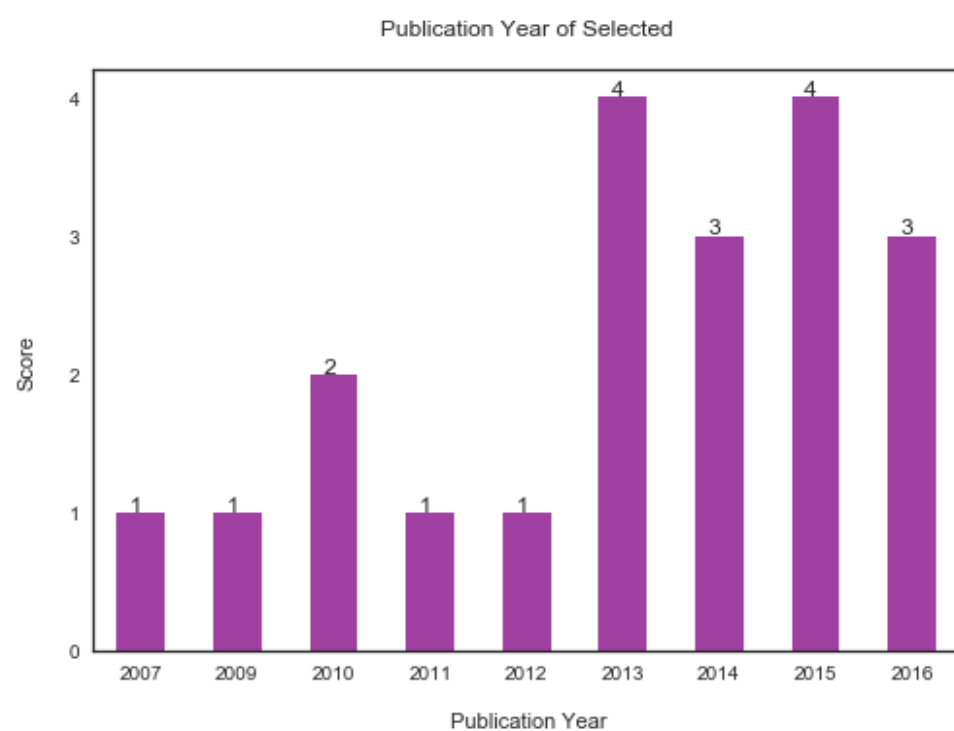
Out[18]:

Digital Library	Visited	Selected	Excel	Excel Selected	Duplicated
acm	210	5	210	5	0
elcompendex	8	2	1	0	7
sciencedirect	249	5	245	4	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



## Bar Chart Publications Year - Strategy 1



## Report Strategy 1 - Seed Set of each Digital Library - Intersection

[illegible]

```

'elcompendex & scopus & springer & webofscience & wiley': [],
'elcompendex & scopus & springer & wiley': [],
'elcompendex & scopus & webofscience': [],
'elcompendex & scopus & webofscience & wiley': [],
'elcompendex & scopus & wiley': [],
'elcompendex & springer': [],
'elcompendex & springer & webofscience': [],
'elcompendex & springer & webofscience & wiley': [],
'elcompendex & springer & wiley': [],
'elcompendex & webofscience': [],
'elcompendex & webofscience & wiley': [],
'elcompendex & wiley': [],
'sciencedirect': ['vlietland2016a',
'vlietland2015a',
'wale2015a',
'eloranta2016a'],
'sciencedirect & scopus': ['hoda2011a'],
'sciencedirect & scopus & springer': [],
'sciencedirect & scopus & springer & webofscience': [],
'sciencedirect & scopus & springer & webofscience & wiley': [],
'sciencedirect & scopus & springer & wiley': [],
'sciencedirect & scopus & webofscience': [],
'sciencedirect & scopus & webofscience & wiley': [],
'sciencedirect & scopus & wiley': [],
'sciencedirect & springer': [],
'sciencedirect & springer & webofscience': [],
'sciencedirect & springer & webofscience & wiley': [],
'sciencedirect & springer & wiley': [],
'sciencedirect & webofscience': [],
'sciencedirect & webofscience & wiley': [],
'sciencedirect & wiley': [],
'scopus': ['rindell2015a', 'ozkaya2013a', 'davis2013a', 'heck2015a'],
'scopus & springer': [],
'scopus & springer & webofscience': [],
'scopus & springer & webofscience & wiley': [],
'scopus & springer & wiley': [],
'scopus & webofscience': [],
'scopus & webofscience & wiley': [],
'scopus & wiley': [],
'springer': [],
'springer & webofscience': [],
'springer & webofscience & wiley': [],
'springer & wiley': [],
'webofscience': [],
'webofscience & wiley': [],
'wiley': []}

```

#### Report Strategy 1 - Seed Set

Out[22]: 16

#### Report Strategy 1 - Unique in the intersection

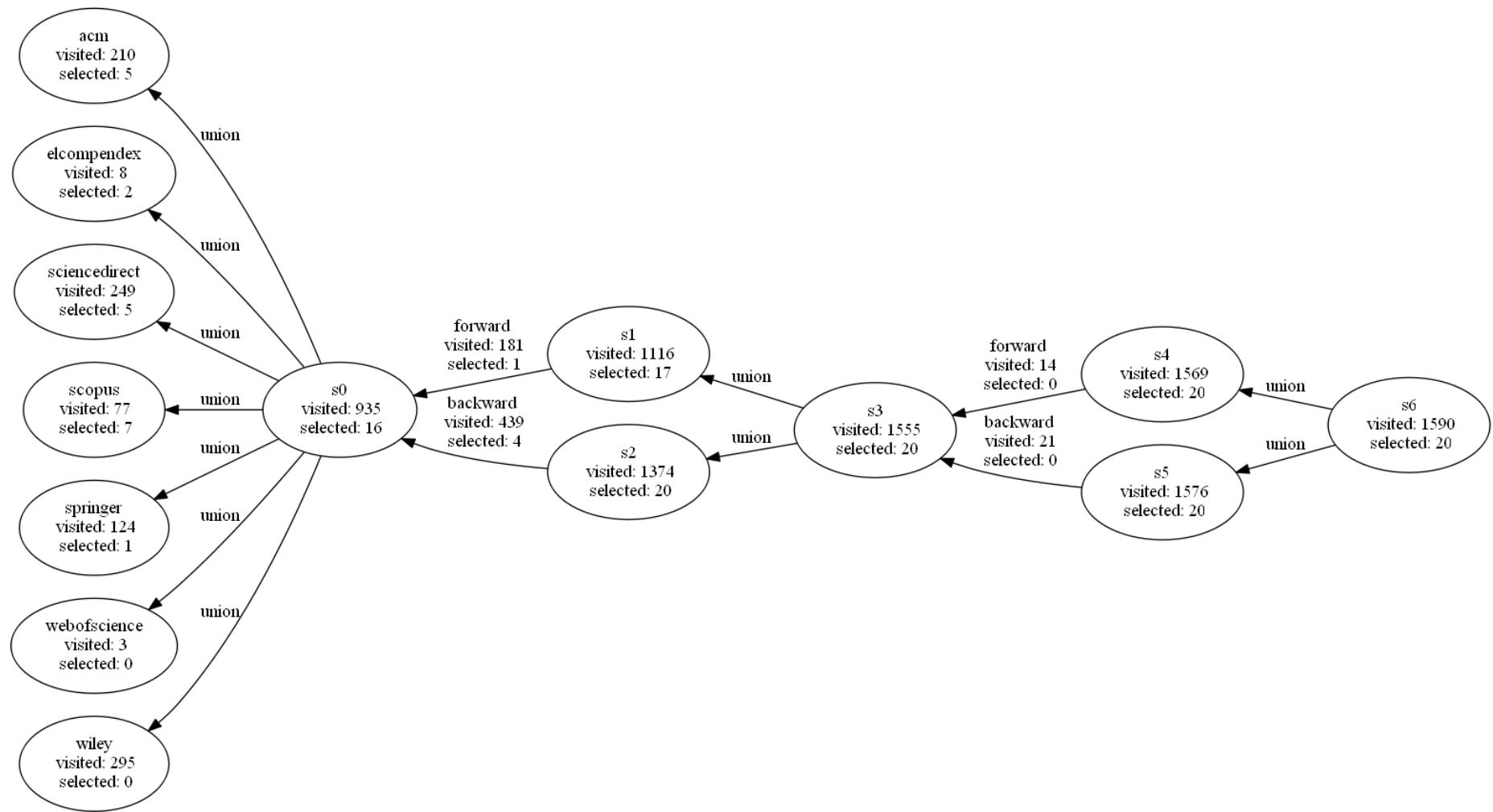
Out[23]: 13

#### Report Strategy 1 - Empty set

Out[24]: ['elcompendex', 'springer', 'webofscience', 'wiley']

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

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



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[27]:

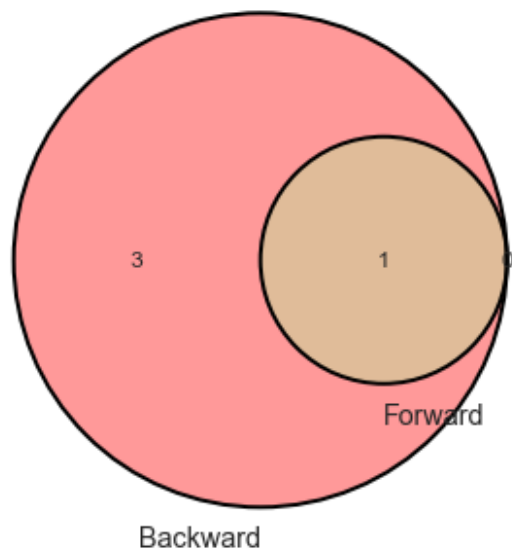
Digital Library/State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
acm	210	5	2.38	25.00	4.35	-	210	5	2.38	25.00	4.35
elcompindex	8	2	25.00	10.00	14.29	-	8	2	25.00	10.00	14.29
sciencedirect	249	5	2.01	25.00	3.72	-	249	5	2.01	25.00	3.72
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	16	1.71	80.00	3.35
s1	181	1	0.55	5.00	1.00	forward	1116	17	1.52	85.00	2.99
s2	439	4	0.91	20.00	1.74	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[28]:

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

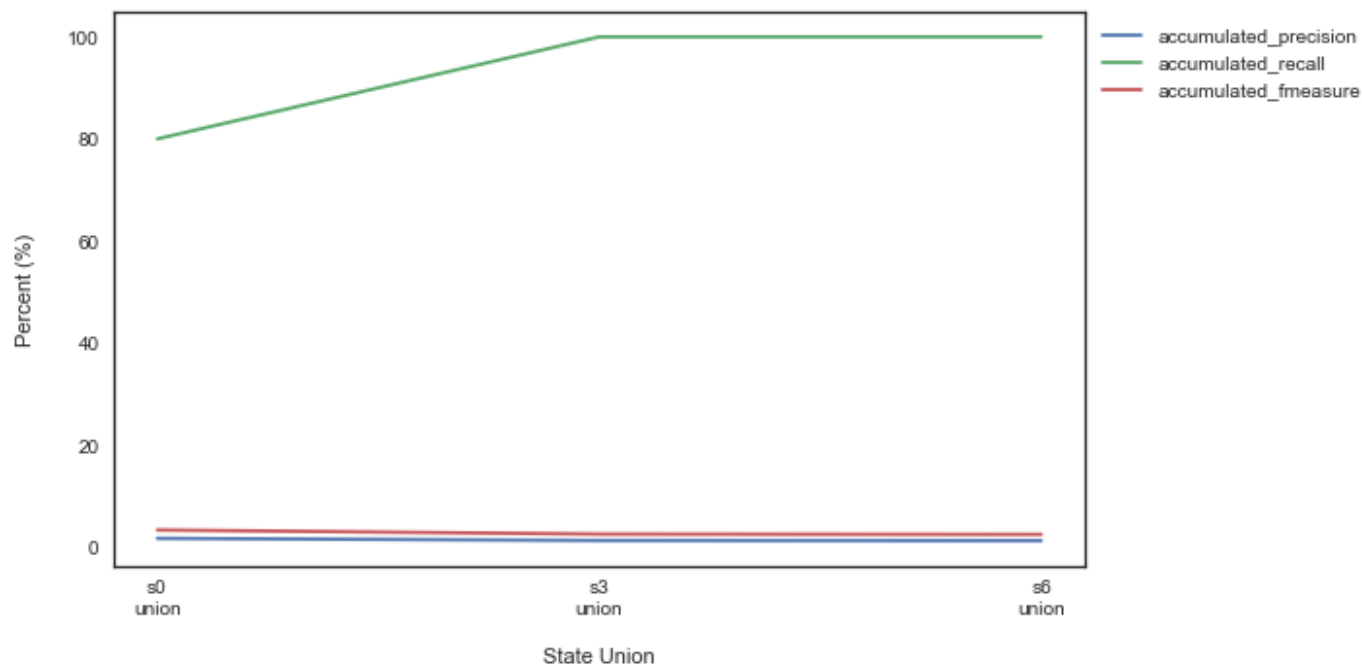
Diagrama de Venn - Strategy 2



Report Diagrama de Venn

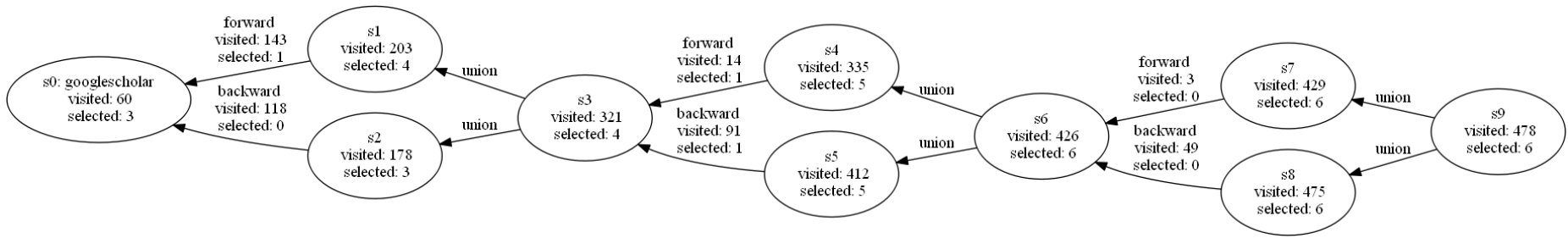
Backward ['barton2009a', 'saddington2012a', 'eloranta2013a', 'waardenburg2013a']  
Forward ['waardenburg2013a']

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[34]:

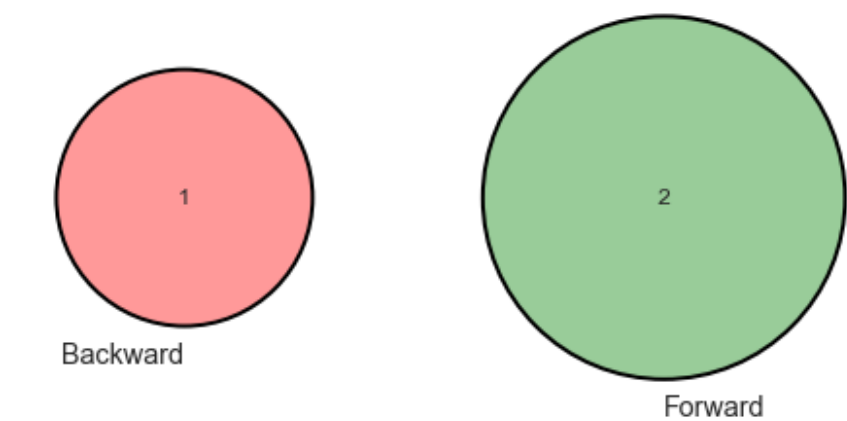
Scholar=s0 /State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
s0: googlescholar	60	3	5.00	15.00	7.50	-	60	3	5.00	15.00	7.50
s1	143	1	0.70	5.00	1.23	forward	203	4	1.97	20.00	3.59
s2	118	0	0.00	0.00	0.00	backward	178	3	1.69	15.00	3.03
s3	0	0	nan	0.00	0.00	union	321	4	1.25	20.00	2.35
s4	14	1	7.14	5.00	5.88	forward	335	5	1.49	25.00	2.82
s5	91	1	1.10	5.00	1.80	backward	412	5	1.21	25.00	2.31
s6	0	0	nan	0.00	0.00	union	426	6	1.41	30.00	2.69
s7	3	0	0.00	0.00	0.00	forward	429	6	1.40	30.00	2.67
s8	49	0	0.00	0.00	0.00	backward	475	6	1.26	30.00	2.42
s9	0	0	nan	0.00	0.00	union	478	6	1.26	30.00	2.41

Report Strategy 3 - Measures of Strategy (accumulated)

Out[35]:

Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
GS + SB	478	6	1.26	30.00	2.41

Diagrama de Venn - Strategy 3

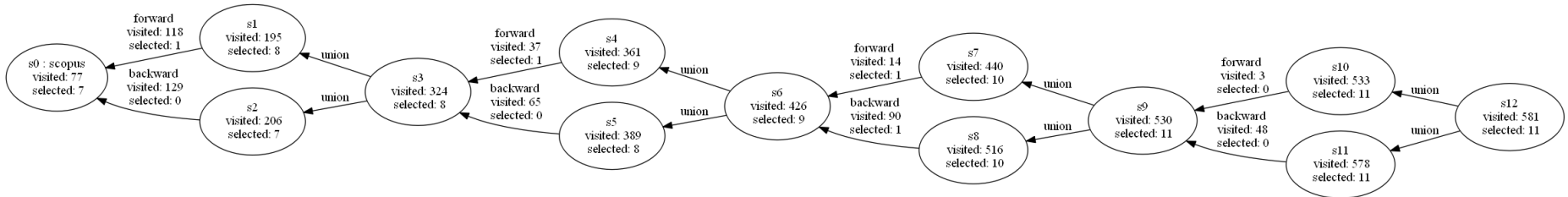


Report Diagrama de Venn

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

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

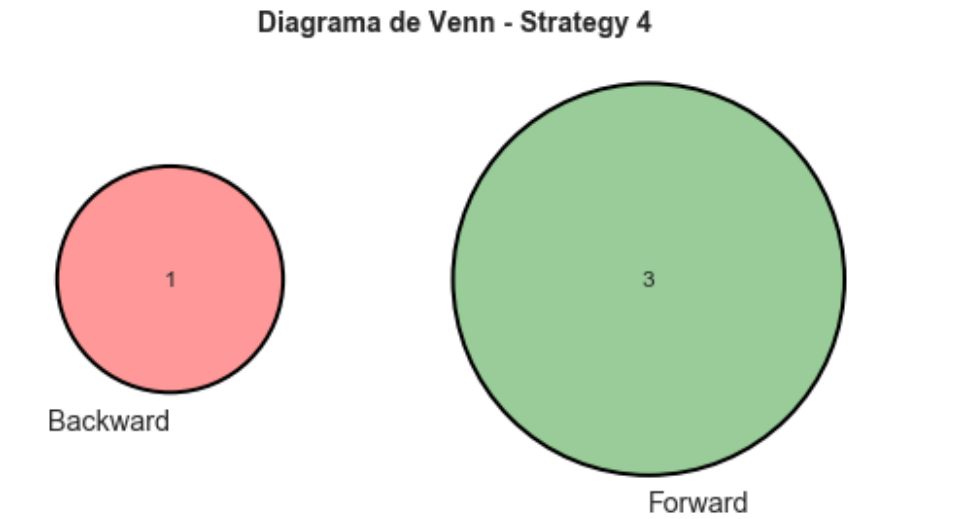
Out[40]:

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[41]:

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

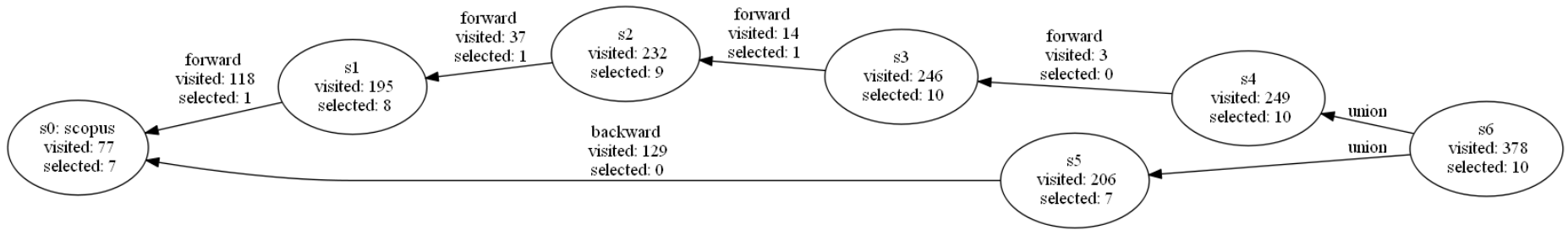


Report Diagrama de Venn

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

▼ 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[46]:

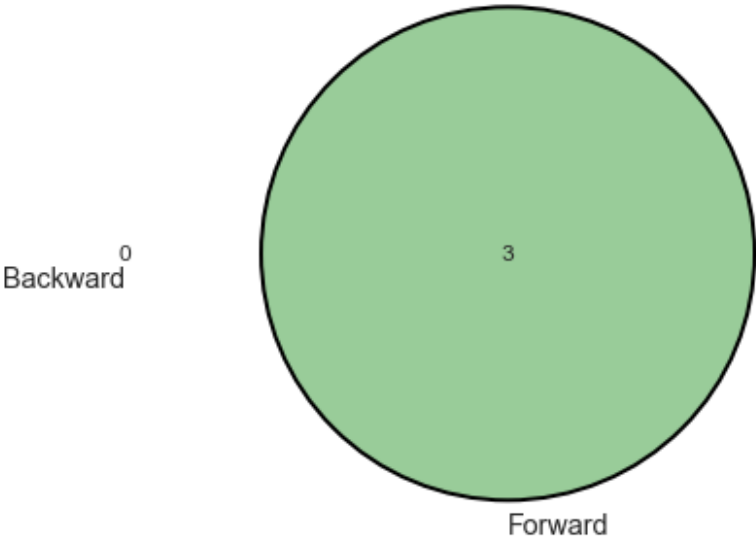
Scopus=s0 /State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
s0: scopus	77	7	9.09	35.00	14.43	-	77	7	9.09	35.00	14.43
s1	118	1	0.85	5.00	1.45	forward	195	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[47]:

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

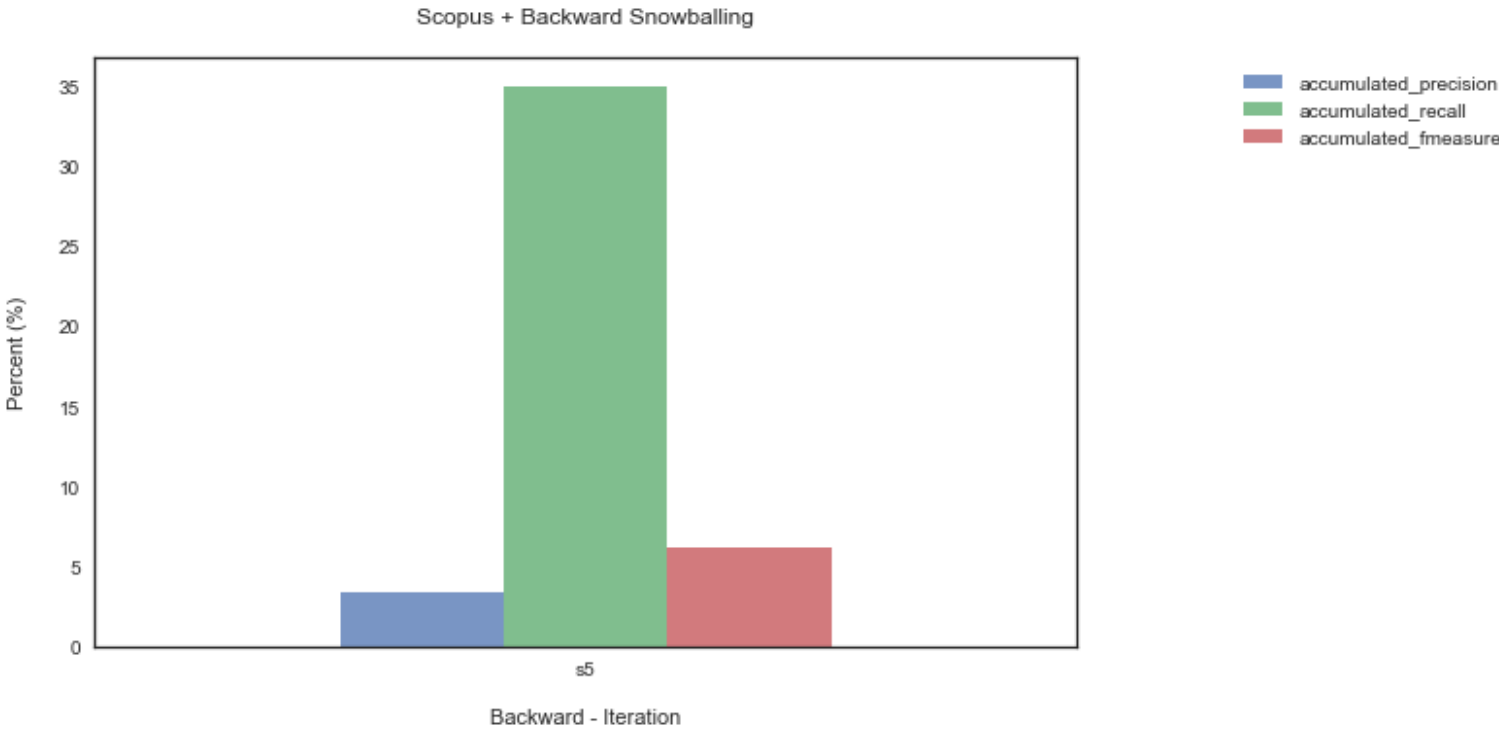
Diagrama de Venn - Strategy 5



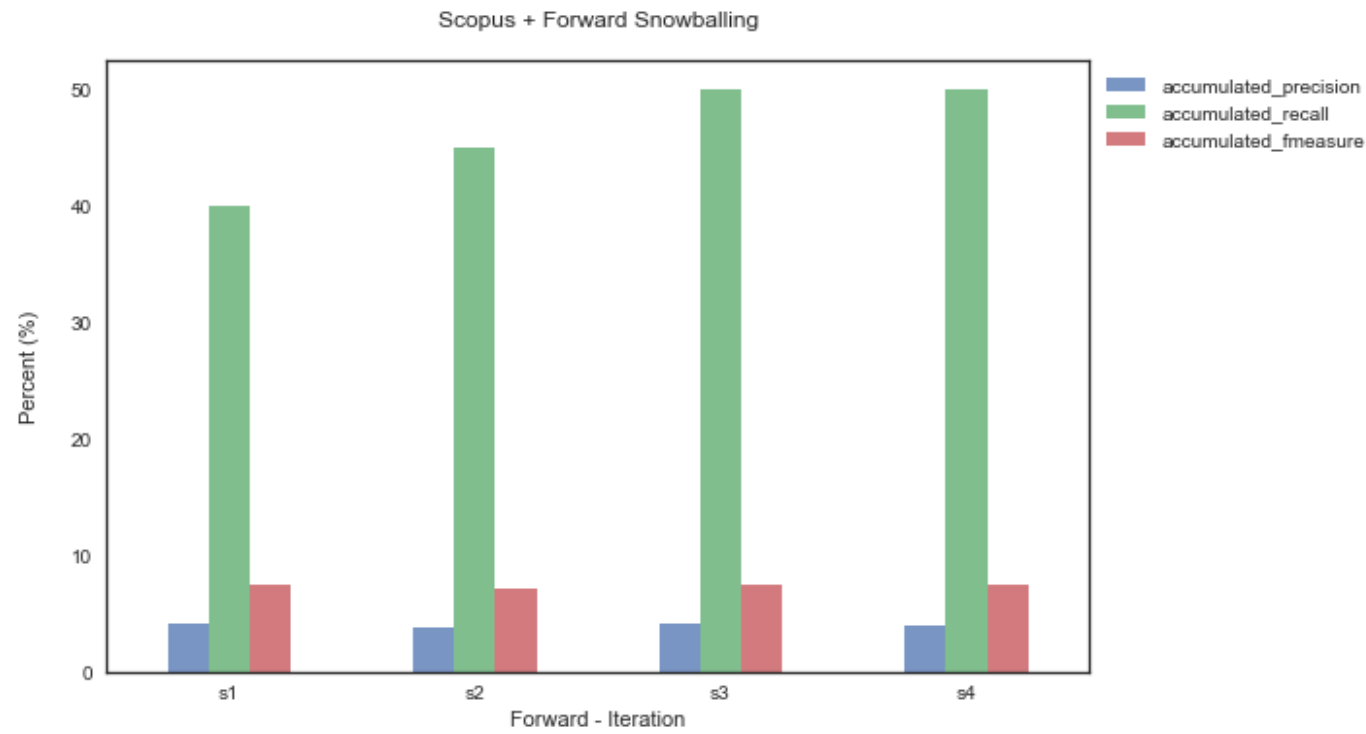
Report Diagrama de Venn

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

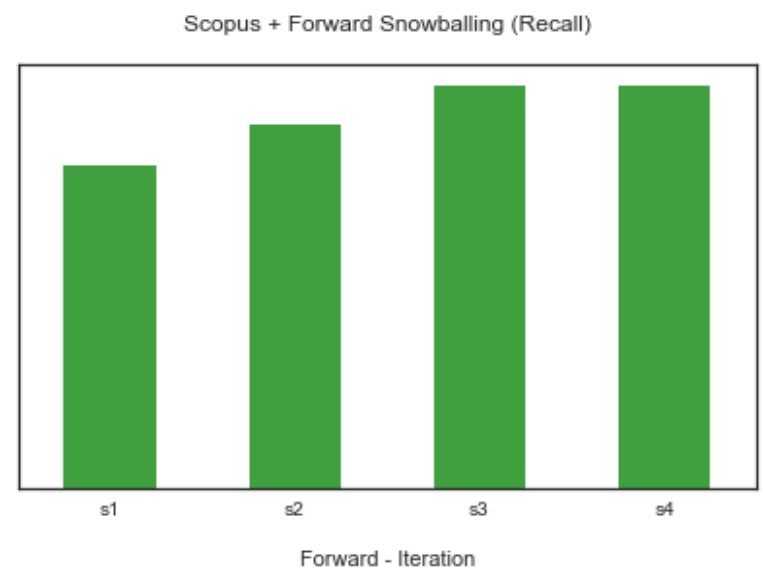
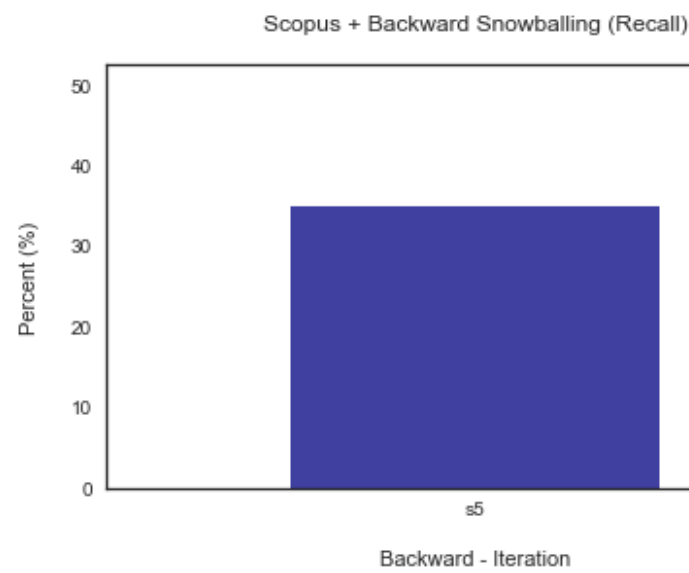
Bar Chart Strategy 5



## Bar Chart Strategy 5

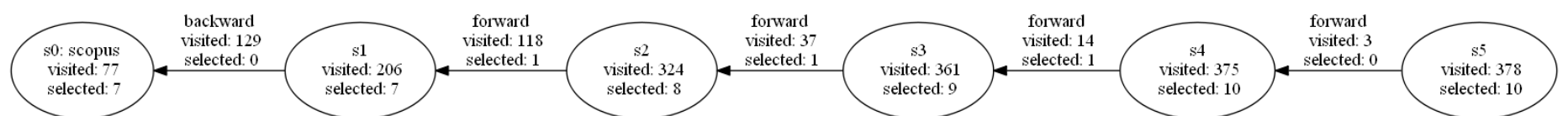


## Bar Chart Strategy 5 - (Scopus + BS // FS)



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

### Provenance Graph - 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[57]:

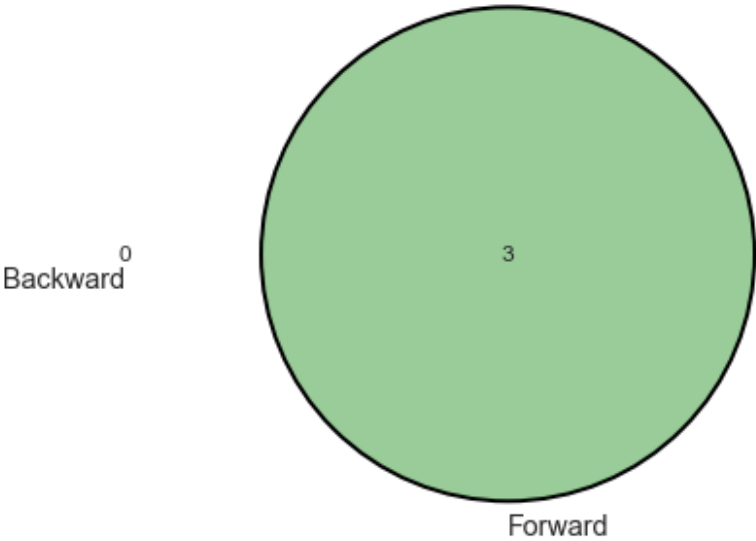
Scopus=s0 /State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
s0: scopus	77	7	9.09	35.00	14.43	-	77	7	9.09	35.00	14.43
s1	129	0	0.00	0.00	0.00	backward	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[58]:

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

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

Provenance Graph - 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[63]:

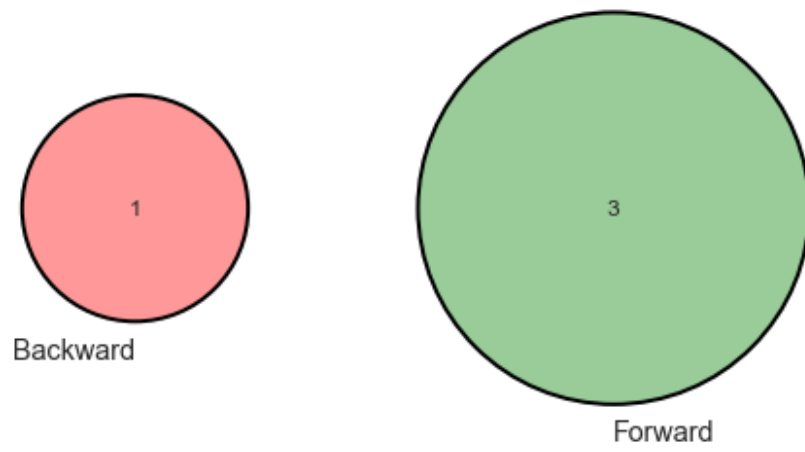
Scopus=s0 /State	Delta Visited	Delta Selected	Delta Precision%	Delta Recall%	Delta F_Measure%	Operation	Accumulated Visited	Accumulated Selected	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%
s0: scopus	77	7	9.09	35.00	14.43	-	77	7	9.09	35.00	14.43
s1	118	1	0.85	5.00	1.45	forward	195	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[64]:

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

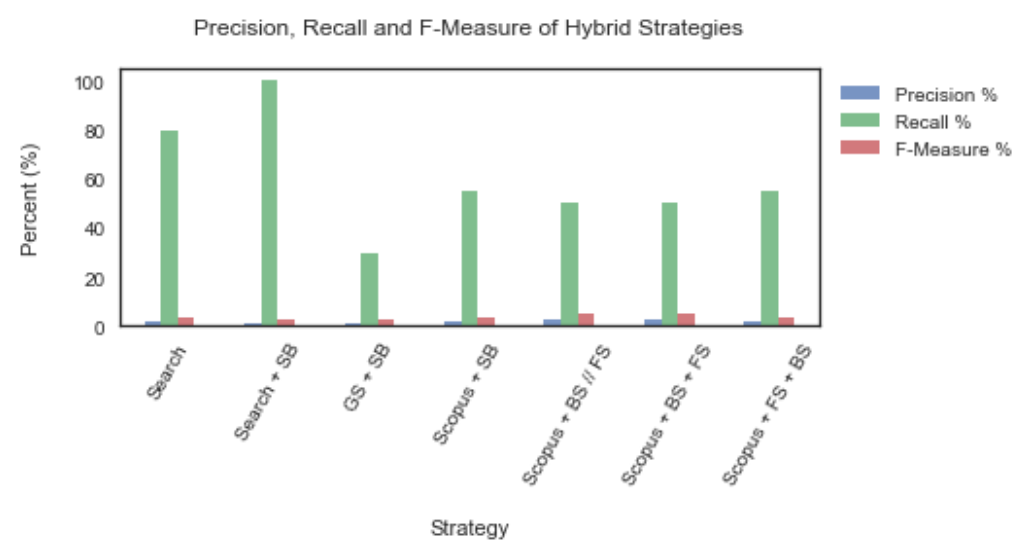
Backward ['saddington2012a']  
 Forward ['vlietland2016a', 'vlietland2015a', 'waardenburg2013a']

## 3 Analyses

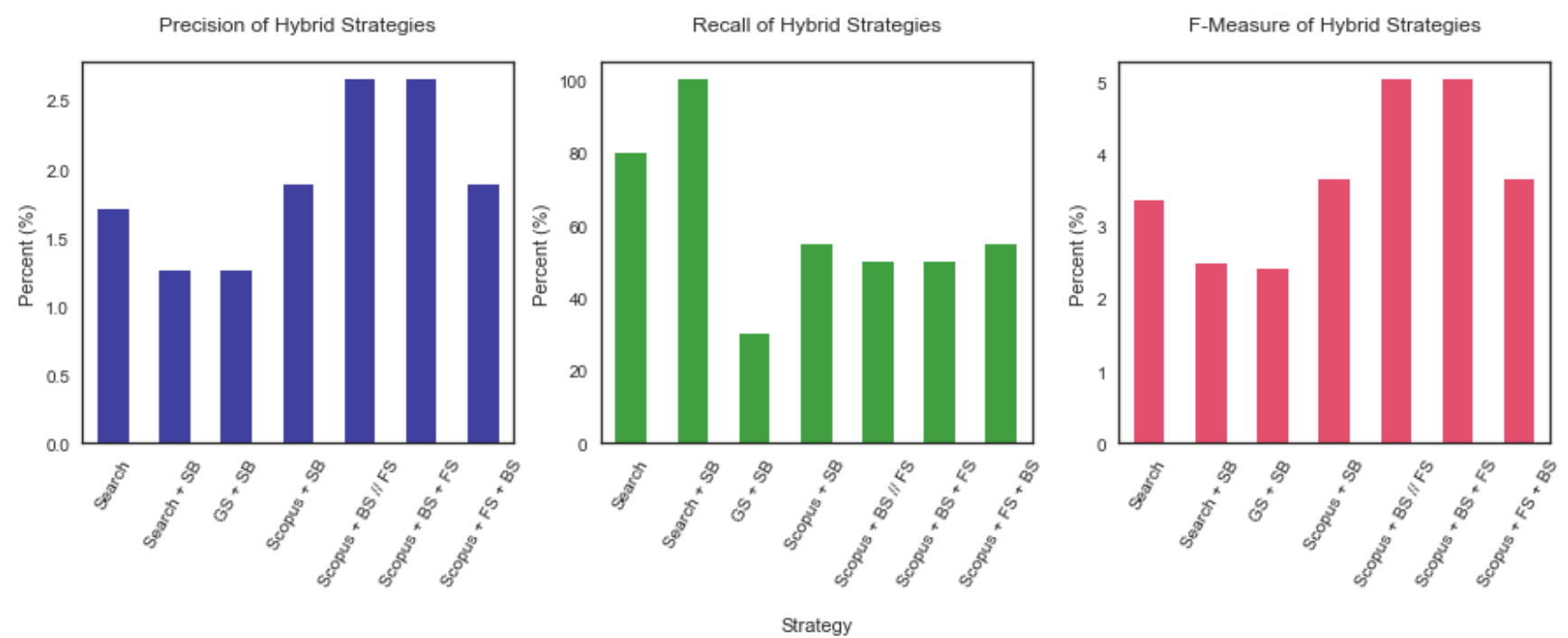
Out[68]:

Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Search	935	16	1.71	80.00	3.35
Search + SB	1590	20	1.26	100.00	2.48
GS + SB	478	6	1.26	30.00	2.41
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



#### Bar Chart Strategies



## ▼ 4 Research Questions

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

Out[76]:

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	5	2.01	25.00	3.72
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[77]:

Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Search + SB	1590	20	1.26	100.00	2.48
GS + SB	478	6	1.26	30.00	2.41
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

RQ3) Which of the strategies is more efficient?

Out[78]:

Strategy	Visited	Selected	Precision %	Recall %	F-Measure %
Search	935	16	1.71	80.00	3.35
Search + SB	1590	20	1.26	100.00	2.48
GS + SB	478	6	1.26	30.00	2.41
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

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

Total of selected studies in the SLR = 20

Out[79]:

Digital Library	Selected	Recall %
acm	5	25.00
elcompendex	19	95.00
sciencedirect	6	30.00
scopus	19	95.00
springer	2	10.00
webofscience	11	55.00
wiley	0	0.00