

▼ 1 Investigating Hybrid Strategies for Systematic Literature Review

Experiment 3

Study

On the use of ontologies in software process assessment: a systematic literature review

Tarhan, Ayça and Giray, Gorkem

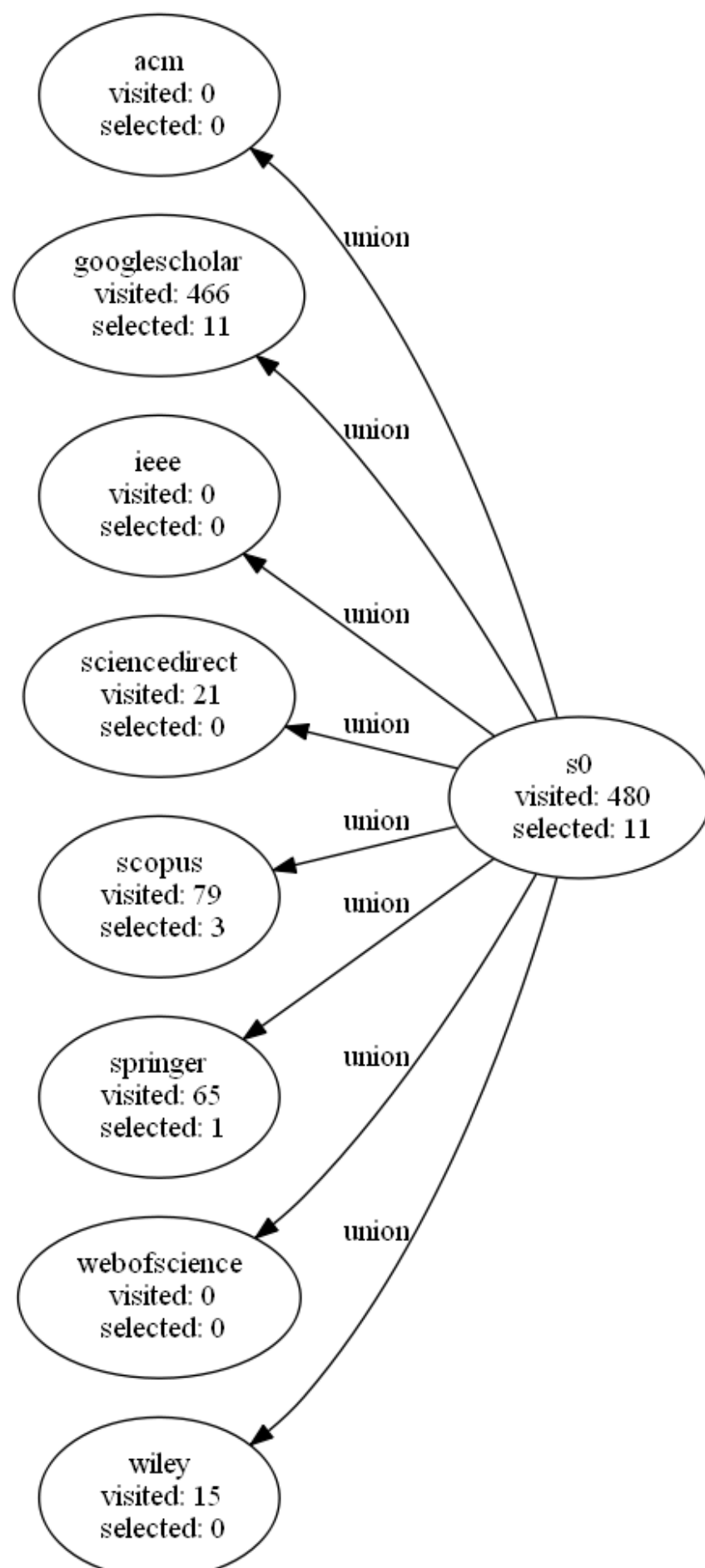
21st International Conference on Evaluation and Assessment in Software Engineering, EASE 2017

▼ 2 Strategies

▼ 2.1 Strategy 1 - Database Search (Search)

Total of selected studies = 11 (Seed Set) + 3 (Snowballing) = 14

Provenance Graph - Strategy 1 - Database Search (Search)



Report Strategy 1 - Database Search (Search)

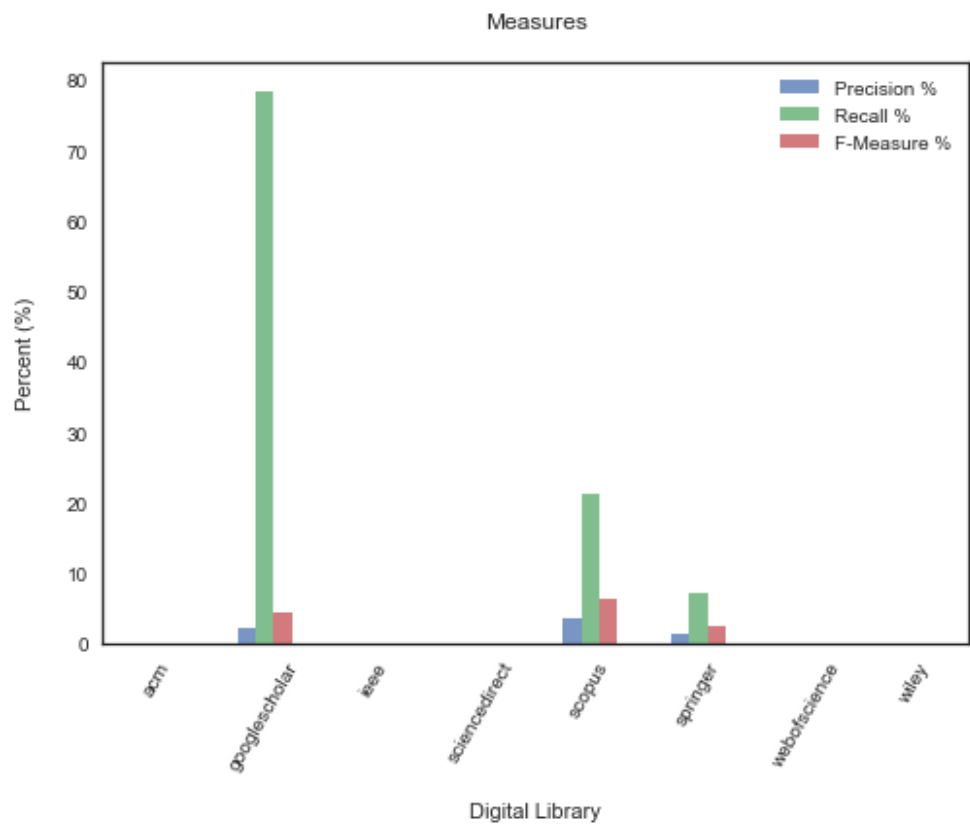
Total of selected studies in the SLR = 14

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

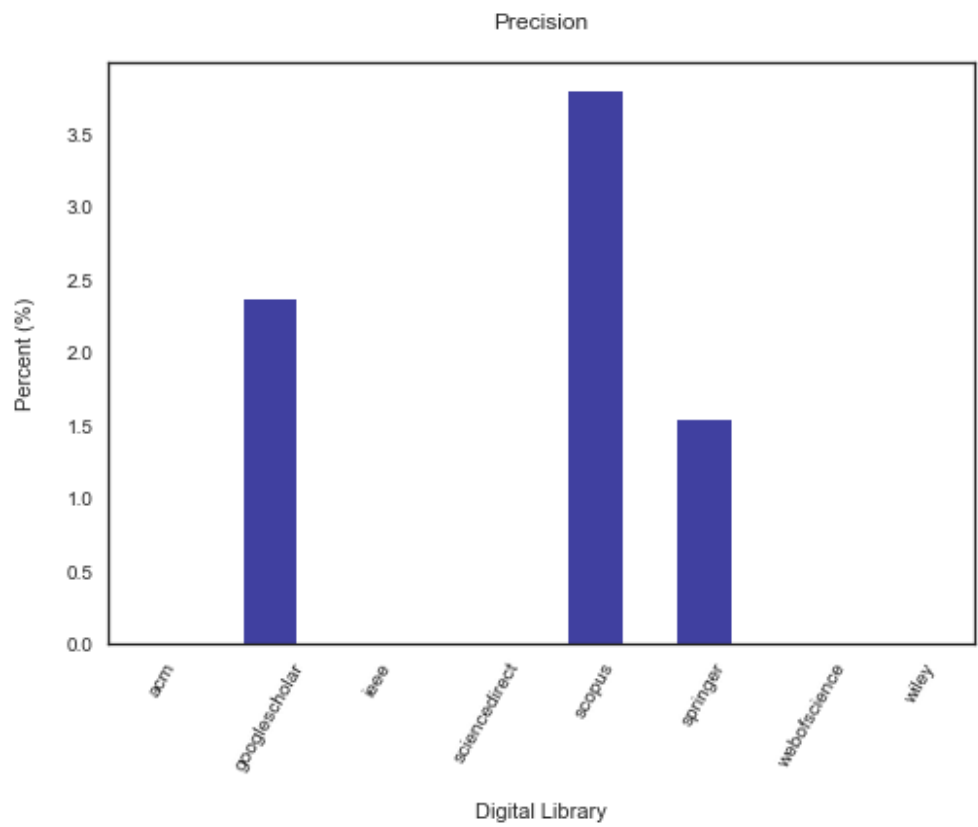
Out[9]:

| Digital Library | Visited | Selected | Precision % | Recall % | F-Measure % |
|-----------------|---------|----------|-------------|----------|-------------|
| acm | 0 | 0 | nan | 0.00 | 0.00 |
| googlescholar | 466 | 11 | 2.36 | 78.57 | 4.58 |
| ieee | 0 | 0 | nan | 0.00 | 0.00 |
| sciencedirect | 21 | 0 | 0.00 | 0.00 | 0.00 |
| scopus | 79 | 3 | 3.80 | 21.43 | 6.45 |
| springer | 65 | 1 | 1.54 | 7.14 | 2.53 |
| webofscience | 0 | 0 | nan | 0.00 | 0.00 |
| wiley | 15 | 0 | 0.00 | 0.00 | 0.00 |

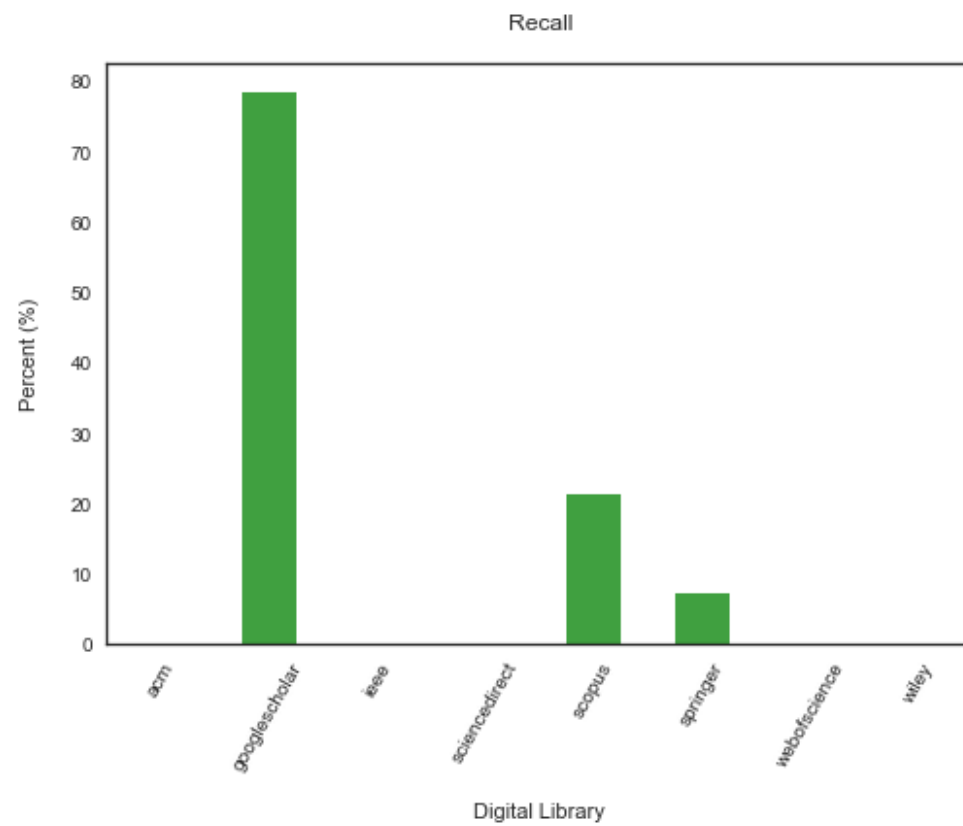
Bar Chart Strategy 1



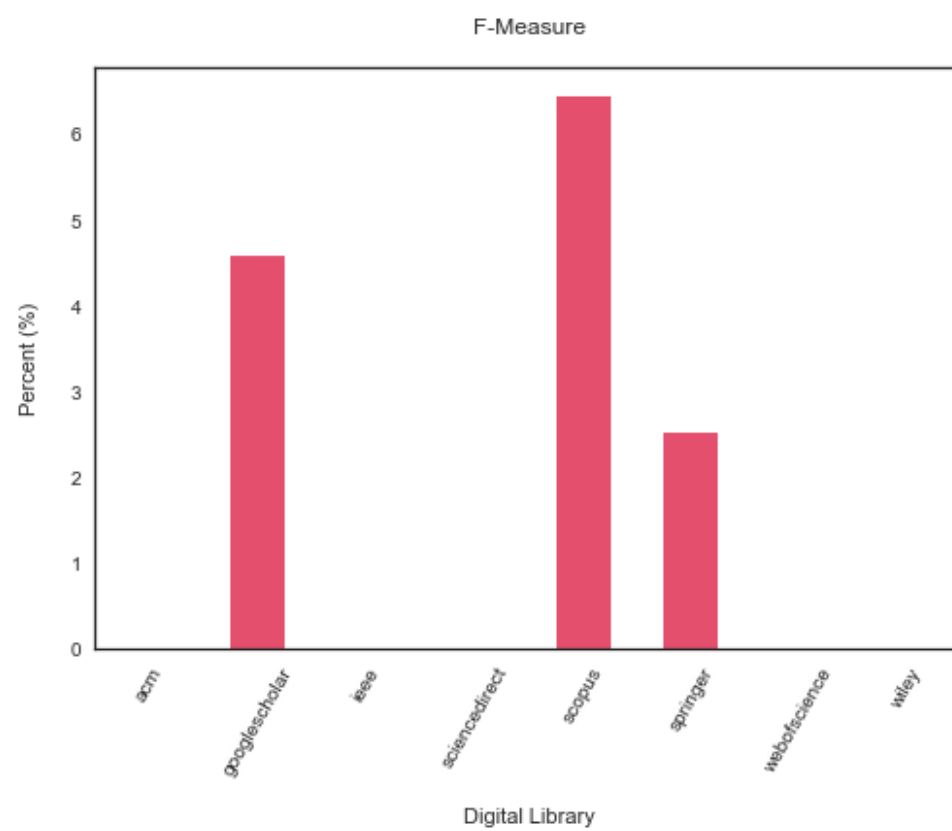
Bar Chart Strategy 1



Bar Chart Strategy 1



Bar Chart Strategy 1



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

Out[15]:

| Strategy | Visited | Selected | Precision % | Recall % | F-Measure % |
|----------|---------|----------|-------------|----------|-------------|
| Search | 480 | 11 | 2.29 | 78.57 | 4.45 |

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

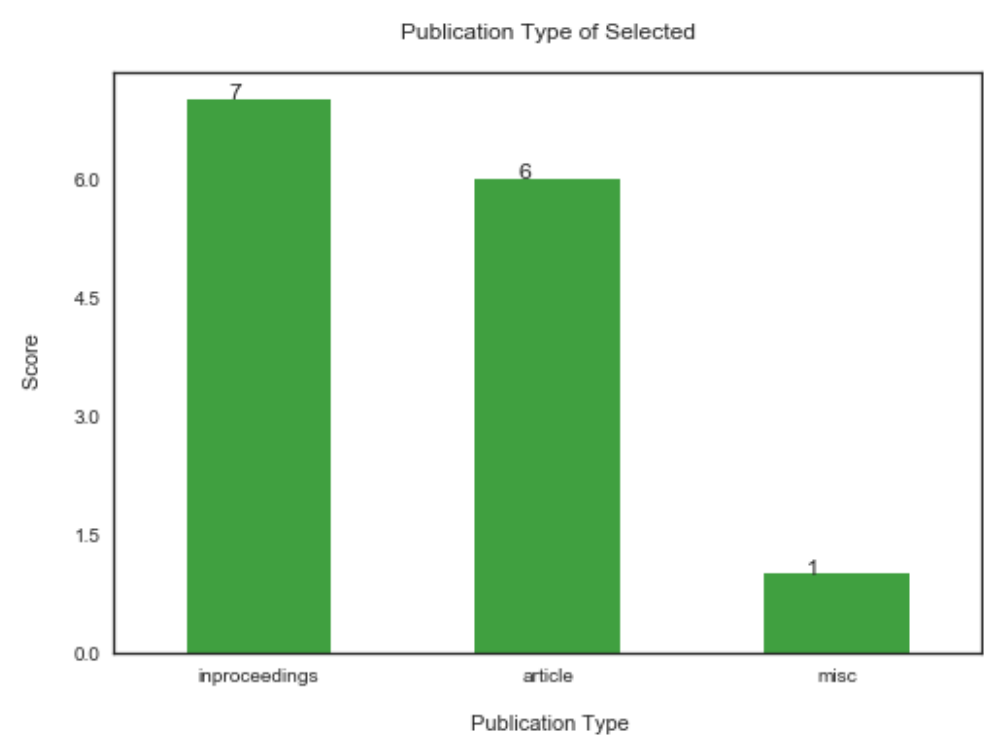
Excel Selected = Papers selected in the SLR and retrieved from Excel provided by authors

Duplicated = Papers identified as duplicate

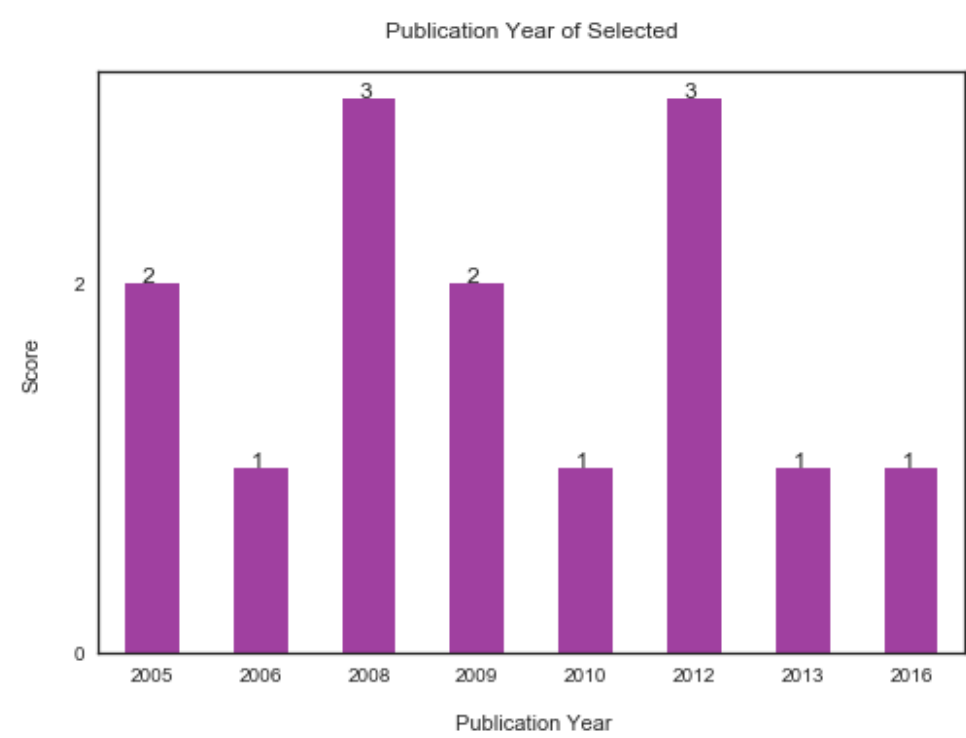
Out[16]:

| Digital Library | Visited | Selected | Excel | Excel Selected | Duplicated |
|-----------------|---------|----------|-------|----------------|------------|
| acm | 0 | 0 | 0 | 0 | 0 |
| googlescholar | 466 | 11 | 466 | 11 | 0 |
| ieee | 0 | 0 | 0 | 0 | 0 |
| sciencedirect | 21 | 0 | 21 | 0 | 0 |
| scopus | 79 | 3 | 2 | 1 | 77 |
| springer | 65 | 1 | 65 | 1 | 0 |
| webofscience | 0 | 0 | 0 | 0 | 0 |
| wiley | 15 | 0 | 15 | 0 | 0 |

Bar Chart Publications Type - Strategy 1

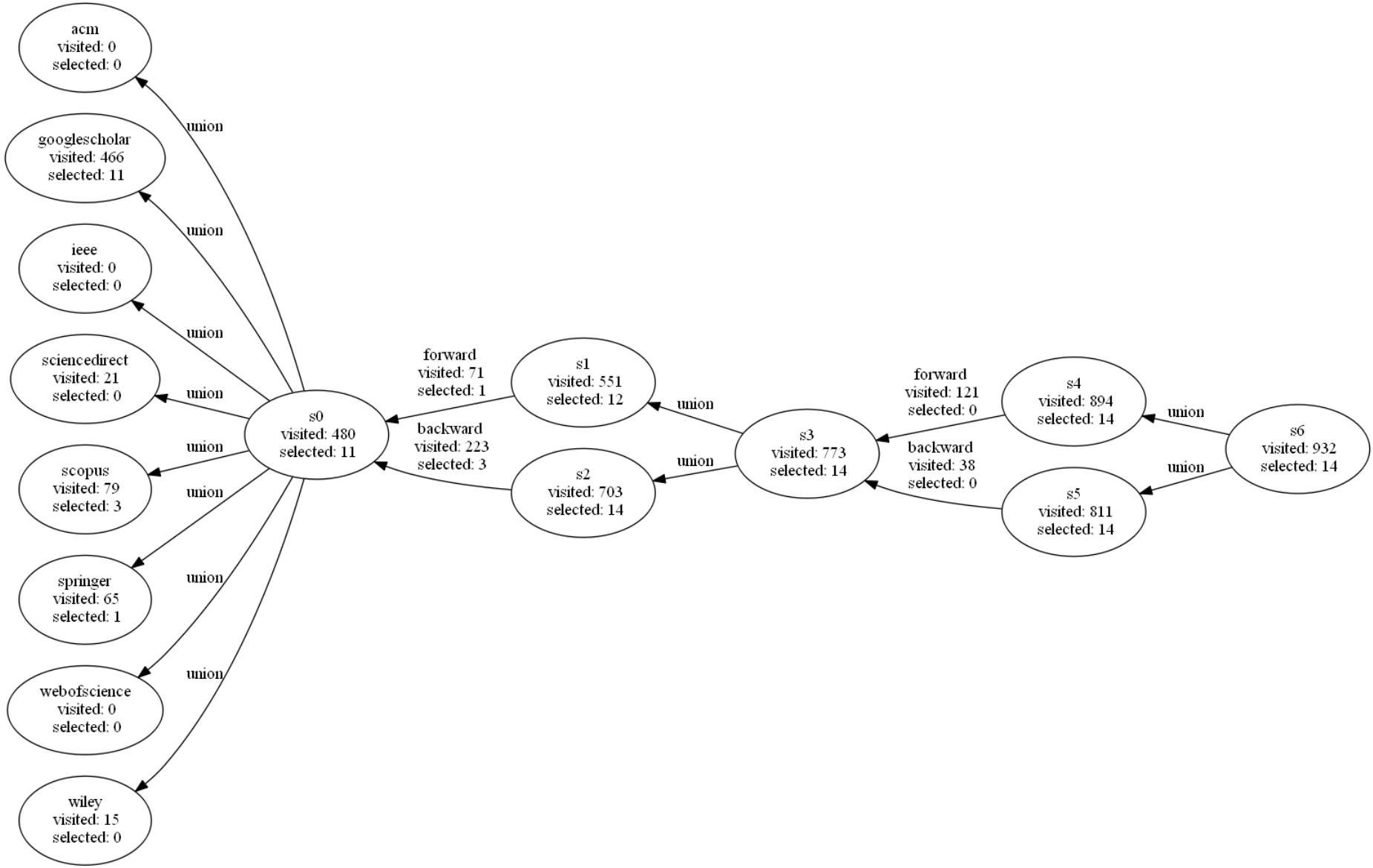


Bar Chart Publications Year - Strategy 1



▼ 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 = 14

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

Out[21]:

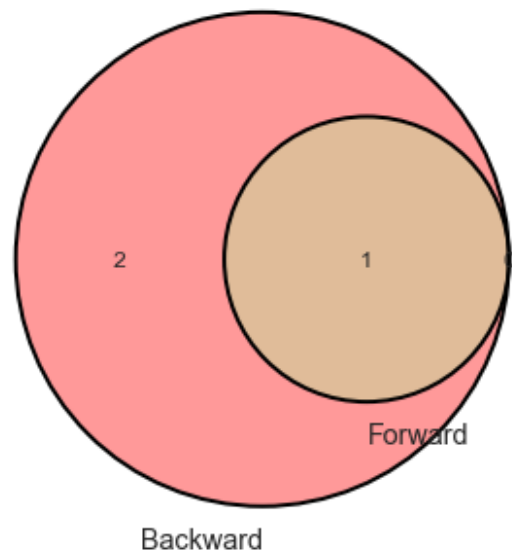
| Digital Library/State | Delta Visited | Delta Selected | Delta Precision% | Delta Recall% | Delta F_Measure% | Operation | Accumulated Visited | Accumulated Selected | Accumulated Precision% | Accumulated Recall% | Accumulated F_Measure% |
|-----------------------|---------------|----------------|------------------|---------------|------------------|-----------|---------------------|----------------------|------------------------|---------------------|------------------------|
| acm | 0 | 0 | nan | 0.00 | 0.00 | - | 0 | 0 | 2.29 | 0.00 | 0.00 |
| googlescholar | 466 | 11 | 2.36 | 78.57 | 4.58 | - | 466 | 11 | 2.36 | 78.57 | 4.58 |
| ieee | 0 | 0 | nan | 0.00 | 0.00 | - | 0 | 0 | 2.36 | 0.00 | 0.00 |
| sciencedirect | 21 | 0 | 0.00 | 0.00 | 0.00 | - | 21 | 0 | 0.00 | 0.00 | 0.00 |
| scopus | 79 | 3 | 3.80 | 21.43 | 6.45 | - | 79 | 3 | 3.80 | 21.43 | 6.45 |
| springer | 65 | 1 | 1.54 | 7.14 | 2.53 | - | 65 | 1 | 1.54 | 7.14 | 2.53 |
| webofscience | 0 | 0 | nan | 0.00 | 0.00 | - | 0 | 0 | 1.54 | 0.00 | 0.00 |
| wiley | 15 | 0 | 0.00 | 0.00 | 0.00 | - | 15 | 0 | 0.00 | 0.00 | 0.00 |
| s0 | 0 | 0 | nan | 0.00 | 0.00 | union | 480 | 11 | 2.29 | 78.57 | 4.45 |
| s1 | 71 | 1 | 1.41 | 7.14 | 2.35 | forward | 551 | 12 | 2.18 | 85.71 | 4.25 |
| s2 | 223 | 3 | 1.35 | 21.43 | 2.53 | backward | 703 | 14 | 1.99 | 100.00 | 3.91 |
| s3 | 0 | 0 | nan | 0.00 | 0.00 | union | 773 | 14 | 1.81 | 100.00 | 3.56 |
| s4 | 121 | 0 | 0.00 | 0.00 | 0.00 | forward | 894 | 14 | 1.57 | 100.00 | 3.08 |
| s5 | 38 | 0 | 0.00 | 0.00 | 0.00 | backward | 811 | 14 | 1.73 | 100.00 | 3.39 |
| s6 | 0 | 0 | nan | 0.00 | 0.00 | union | 932 | 14 | 1.50 | 100.00 | 2.96 |

Report Strategy 2 - Measures of Strategy (accumulated)

Out[22]:

| Strategy | Visited | Selected | Precision % | Recall % | F-Measure % |
|-------------|---------|----------|-------------|----------|-------------|
| Search + SB | 932 | 14 | 1.50 | 100.00 | 2.96 |

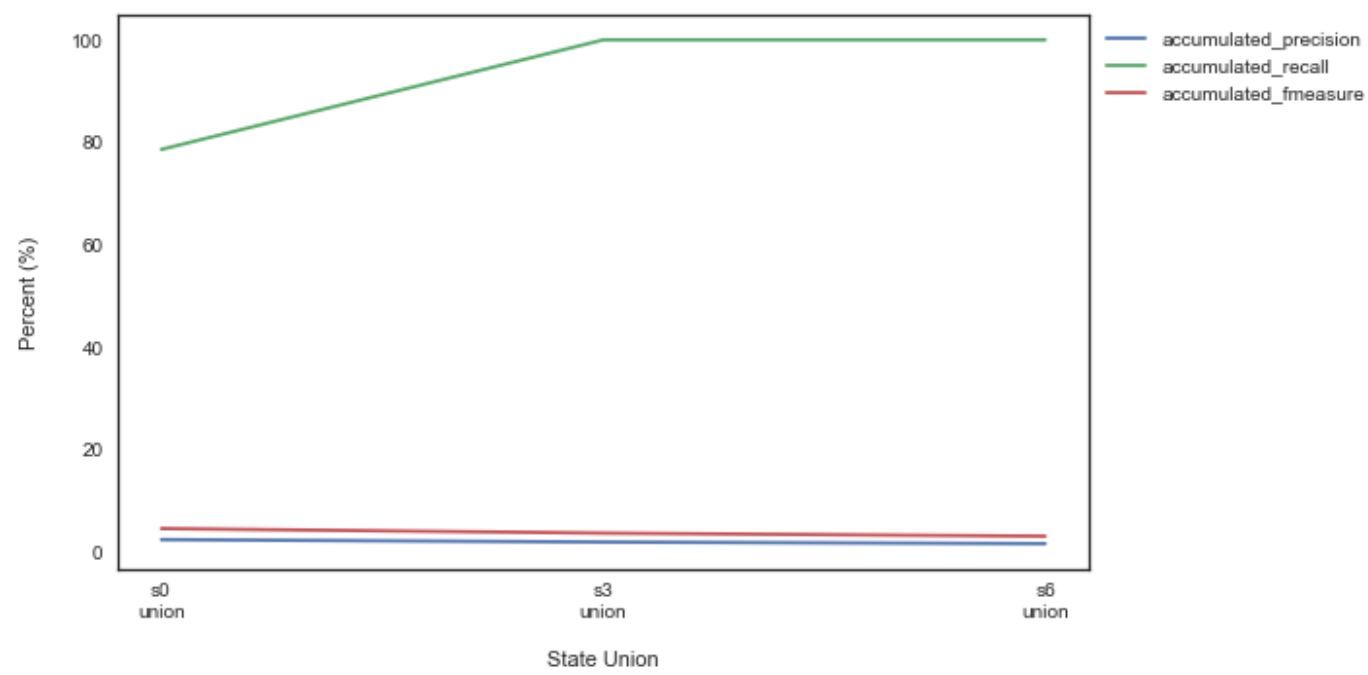
Diagrama de Venn - Strategy 2



Report Diagrama de Venn

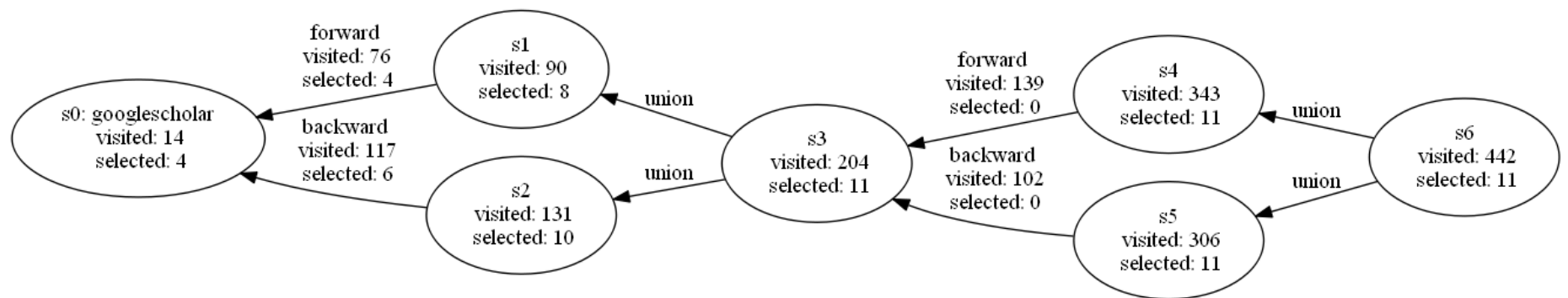
Backward ['soydan2006a', 'lee2008a', 'lee2009a']
Forward ['soydan2006a']

Accumulated Measures - Strategy 2



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

Provenance Graph - Strategy 3 - Informal Search (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 = 14

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

Out[28]:

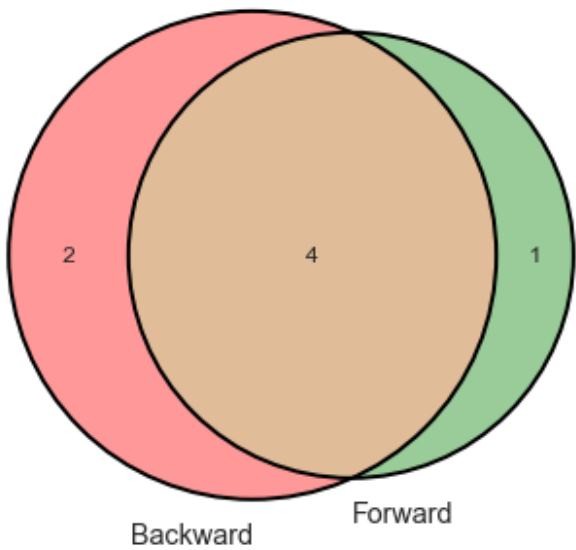
| Scholar=s0 /State | Delta Visited | Delta Selected | Delta Precision% | Delta Recall% | Delta F_Measure% | Operation | Accumulated Visited | Accumulated Selected | Accumulated Precision% | Accumulated Recall% | Accumulated F_Measure% |
|-------------------|---------------|----------------|------------------|---------------|------------------|-----------|---------------------|----------------------|------------------------|---------------------|------------------------|
| s0: googlescholar | 14 | 4 | 28.57 | 28.57 | 28.57 | - | 14 | 4 | 28.57 | 28.57 | 28.57 |
| s1 | 76 | 4 | 5.26 | 28.57 | 8.89 | forward | 90 | 8 | 8.89 | 57.14 | 15.38 |
| s2 | 117 | 6 | 5.13 | 42.86 | 9.16 | backward | 131 | 10 | 7.63 | 71.43 | 13.79 |
| s3 | 0 | 0 | nan | 0.00 | 0.00 | union | 204 | 11 | 5.39 | 78.57 | 10.09 |
| s4 | 139 | 0 | 0.00 | 0.00 | 0.00 | forward | 343 | 11 | 3.21 | 78.57 | 6.16 |
| s5 | 102 | 0 | 0.00 | 0.00 | 0.00 | backward | 306 | 11 | 3.59 | 78.57 | 6.87 |
| s6 | 0 | 0 | nan | 0.00 | 0.00 | union | 442 | 11 | 2.49 | 78.57 | 4.82 |

Report Strategy 3 - Measures of Strategy (accumulated)

Out[29]:

| Strategy | Visited | Selected | Precision % | Recall % | F-Measure % |
|----------|---------|----------|-------------|----------|-------------|
| GS + SB | 442 | 11 | 2.49 | 78.57 | 4.82 |

Diagrama de Venn - Strategy 3

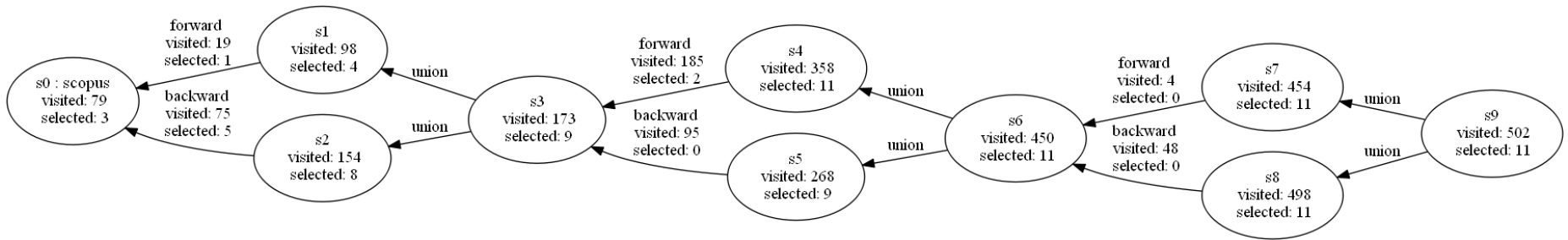


Report Diagrama de Venn

Backward ['grambow2012a', 'soydan2006a', 'rungratri2008a', 'gazel2009a', 'lee2008a', 'lee2009a']
Forward ['grambow2012a', 'rungratri2008a', 'soydan2006a', 'gazel2009a', 'soydan2012a']

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

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



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

Total of selected studies in the SLR = 14

Precision = Selected/ Visited

Recall = Selected/ Total of selected studies in the SLR

F-Measure = Harmonic Mean of Precision and Recall

Out[34]:

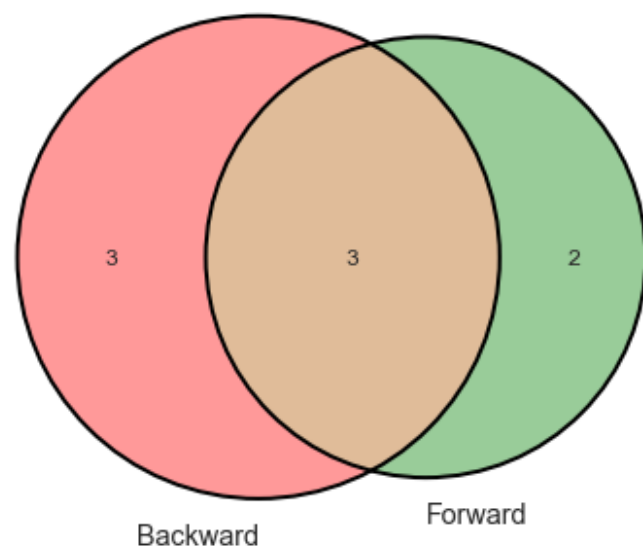
| Scopus=s0 /State | Delta Visited | Delta Selected | Delta Precision% | Delta Recall% | Delta F_Measure% | Operation | Accumulated Visited | Accumulated Selected | Accumulated Precision% | Accumulated Recall% | Accumulated F_Measure% |
|------------------|---------------|----------------|------------------|---------------|------------------|-----------|---------------------|----------------------|------------------------|---------------------|------------------------|
| s0 : scopus | 79 | 3 | 3.80 | 21.43 | 6.45 | - | 79 | 3 | 3.80 | 21.43 | 6.45 |
| s1 | 19 | 1 | 5.26 | 7.14 | 6.06 | forward | 98 | 4 | 4.08 | 28.57 | 7.14 |
| s2 | 75 | 5 | 6.67 | 35.71 | 11.24 | backward | 154 | 8 | 5.19 | 57.14 | 9.52 |
| s3 | 0 | 0 | nan | 0.00 | 0.00 | union | 173 | 9 | 5.20 | 64.29 | 9.63 |
| s4 | 185 | 2 | 1.08 | 14.29 | 2.01 | forward | 358 | 11 | 3.07 | 78.57 | 5.91 |
| s5 | 95 | 0 | 0.00 | 0.00 | 0.00 | backward | 268 | 9 | 3.36 | 64.29 | 6.38 |
| s6 | 0 | 0 | nan | 0.00 | 0.00 | union | 450 | 11 | 2.44 | 78.57 | 4.74 |
| s7 | 4 | 0 | 0.00 | 0.00 | 0.00 | forward | 454 | 11 | 2.42 | 78.57 | 4.70 |
| s8 | 48 | 0 | 0.00 | 0.00 | 0.00 | backward | 498 | 11 | 2.21 | 78.57 | 4.30 |
| s9 | 0 | 0 | nan | 0.00 | 0.00 | union | 502 | 11 | 2.19 | 78.57 | 4.26 |

Report Strategy 4 - Measures of Strategy (accumulated)

Out[35]:

| Strategy | Visited | Selected | Precision % | Recall % | F-Measure % |
|-------------|---------|----------|-------------|----------|-------------|
| Scopus + SB | 502 | 11 | 2.19 | 78.57 | 4.26 |

Diagrama de Venn - Strategy 4



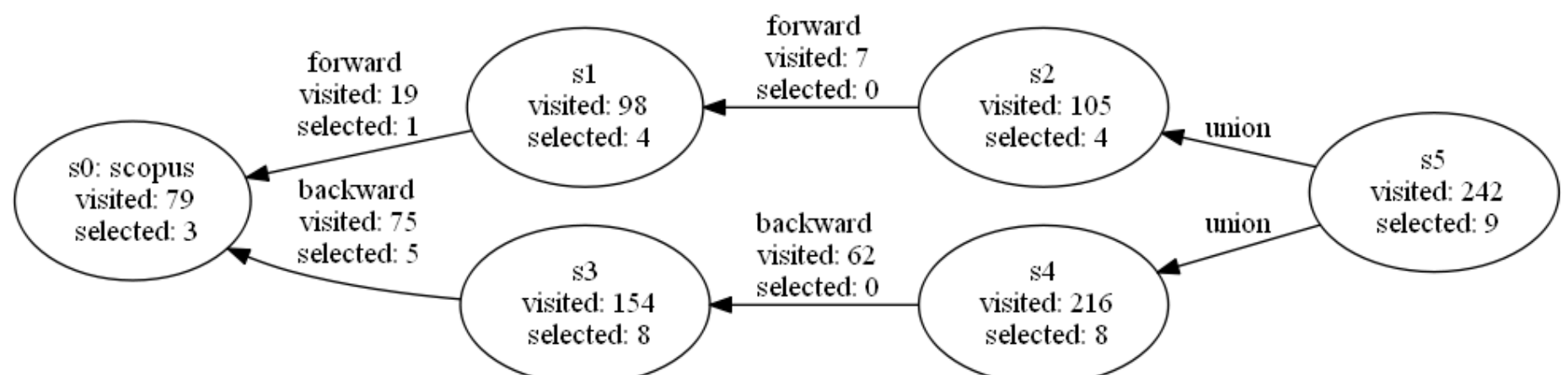
Report Diagrama de Venn

Backward ['grambow2012a', 'soydan2006a', 'rungratri2008a', 'lee2008a', 'lee2009a', 'liao2005a']

Forward ['grambow2013a', 'rungratri2008a', 'grambow2012a', 'soydan2006a', 'soydan2012a']

▼ 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 = 14

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

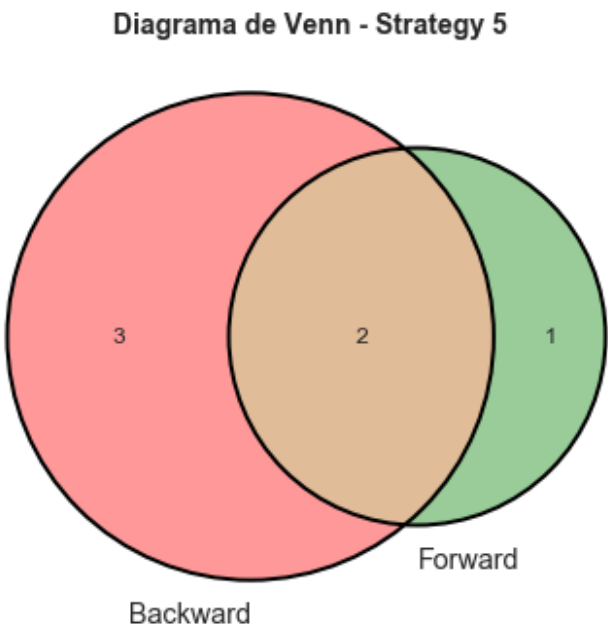
Out[40]:

| 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 | 79 | 3 | 3.80 | 21.43 | 6.45 | - | 79 | 3 | 3.80 | 21.43 | 6.45 |
| s1 | 19 | 1 | 5.26 | 7.14 | 6.06 | forward | 98 | 4 | 4.08 | 28.57 | 7.14 |
| s2 | 7 | 0 | 0.00 | 0.00 | 0.00 | forward | 105 | 4 | 3.81 | 28.57 | 6.72 |
| s3 | 75 | 5 | 6.67 | 35.71 | 11.24 | backward | 154 | 8 | 5.19 | 57.14 | 9.52 |
| s4 | 62 | 0 | 0.00 | 0.00 | 0.00 | backward | 216 | 8 | 3.70 | 57.14 | 6.96 |
| s5 | 0 | 0 | nan | 0.00 | 0.00 | union | 242 | 9 | 3.72 | 64.29 | 7.03 |

Report Strategy 5 - Measures of Strategy (accumulated)

Out[41]:

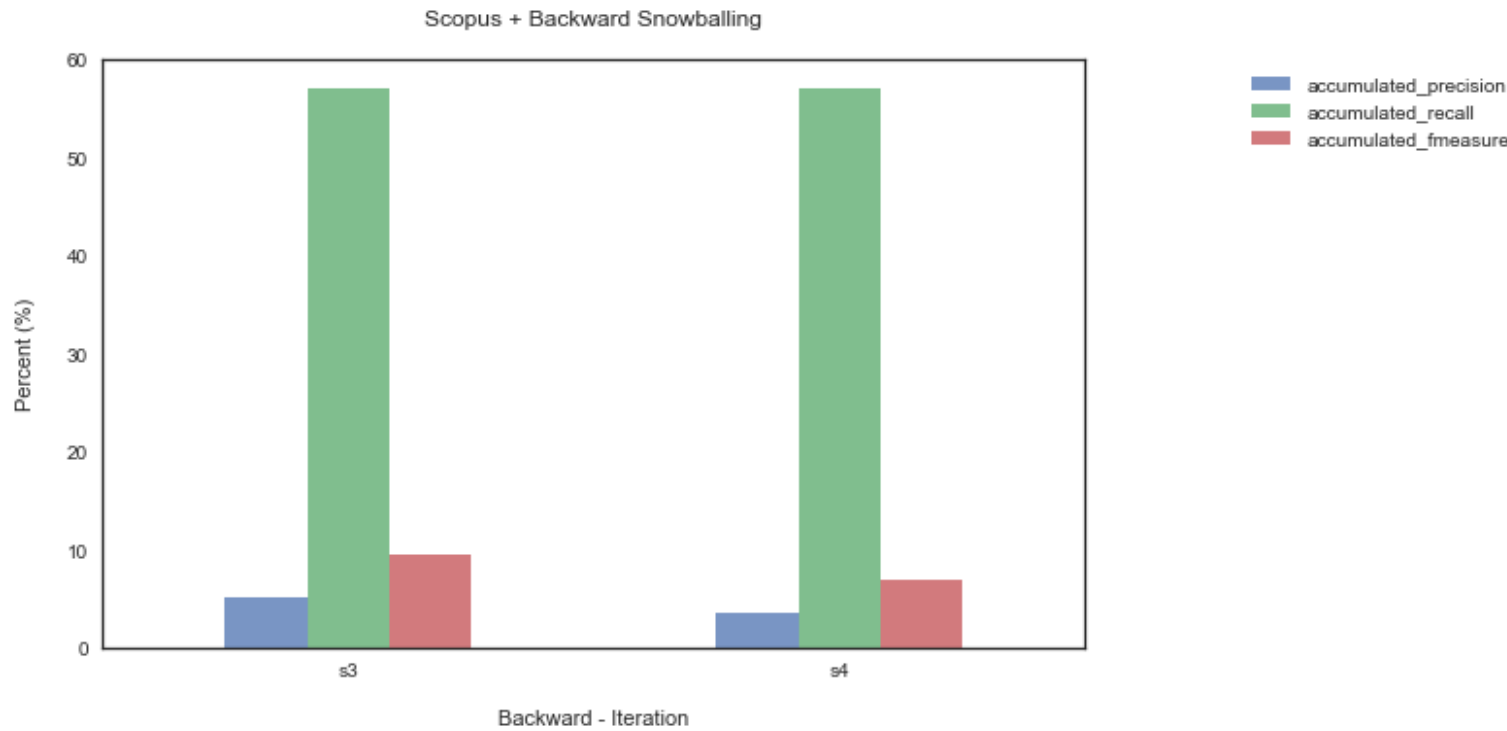
| Strategy | Visited | Selected | Precision % | Recall % | F-Measure % |
|-------------------|---------|----------|-------------|----------|-------------|
| Scopus + BS // FS | 242 | 9 | 3.72 | 64.29 | 7.03 |



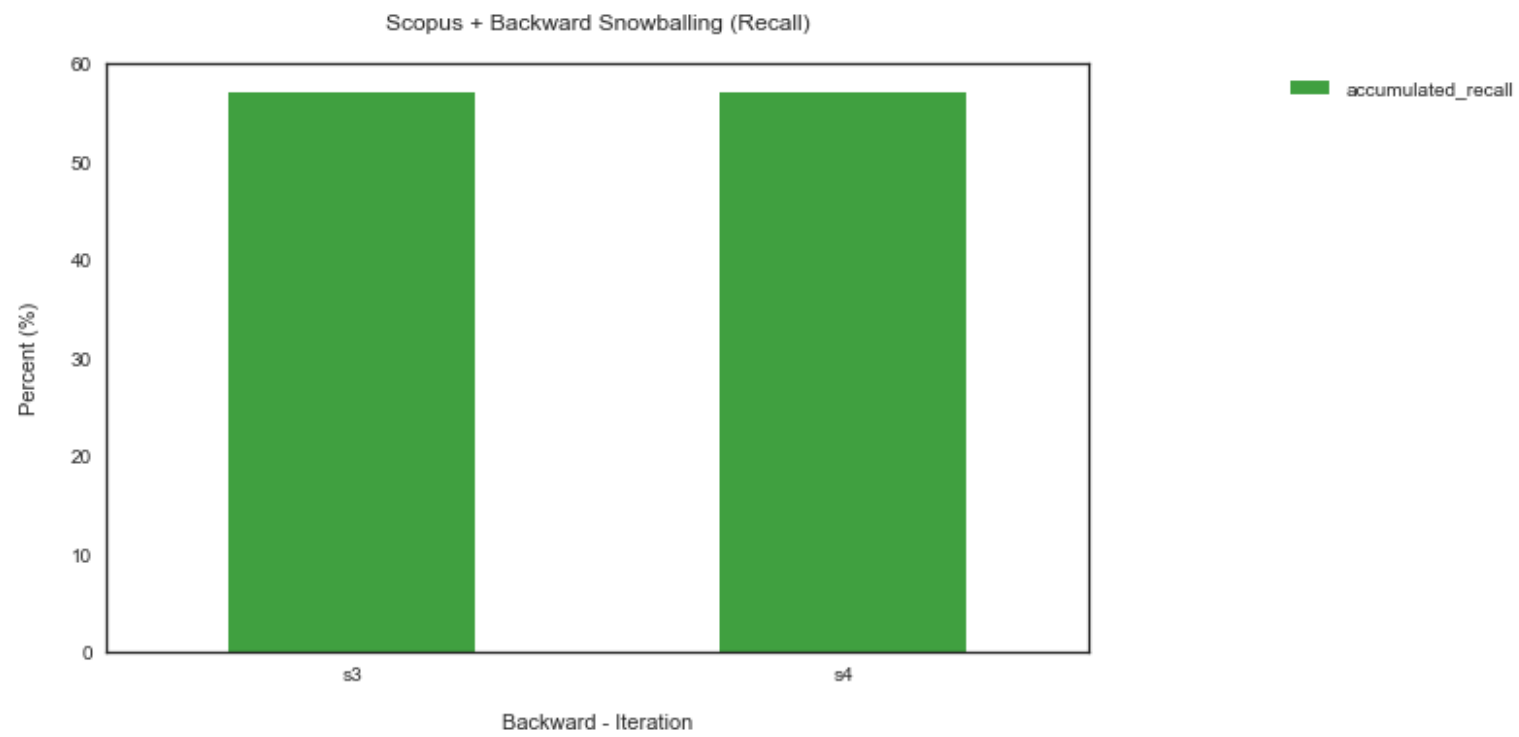
Report Diagrama de Venn

Backward ['soydan2006a', 'rungratri2008a', 'lee2008a', 'liao2005a', 'lee2009a']
Forward ['soydan2012a', 'rungratri2008a', 'soydan2006a']

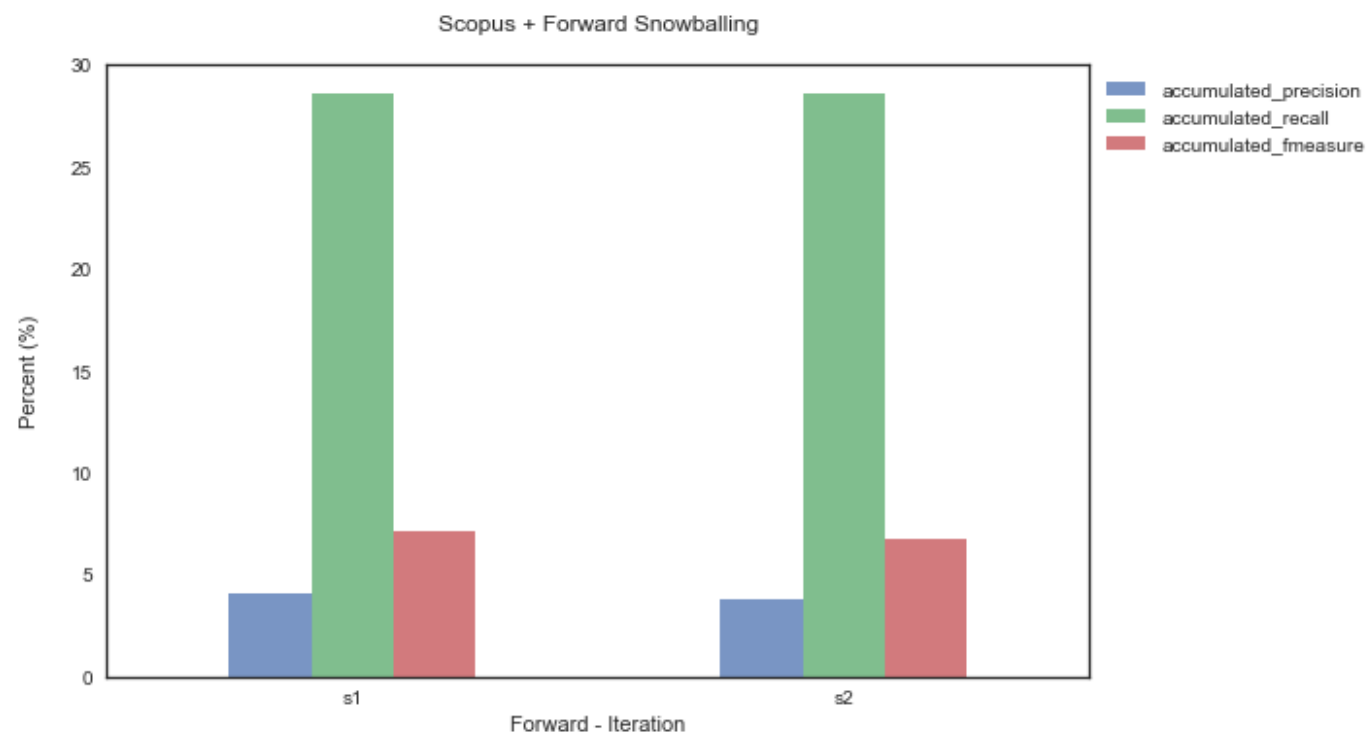
Bar Chart Strategy 5



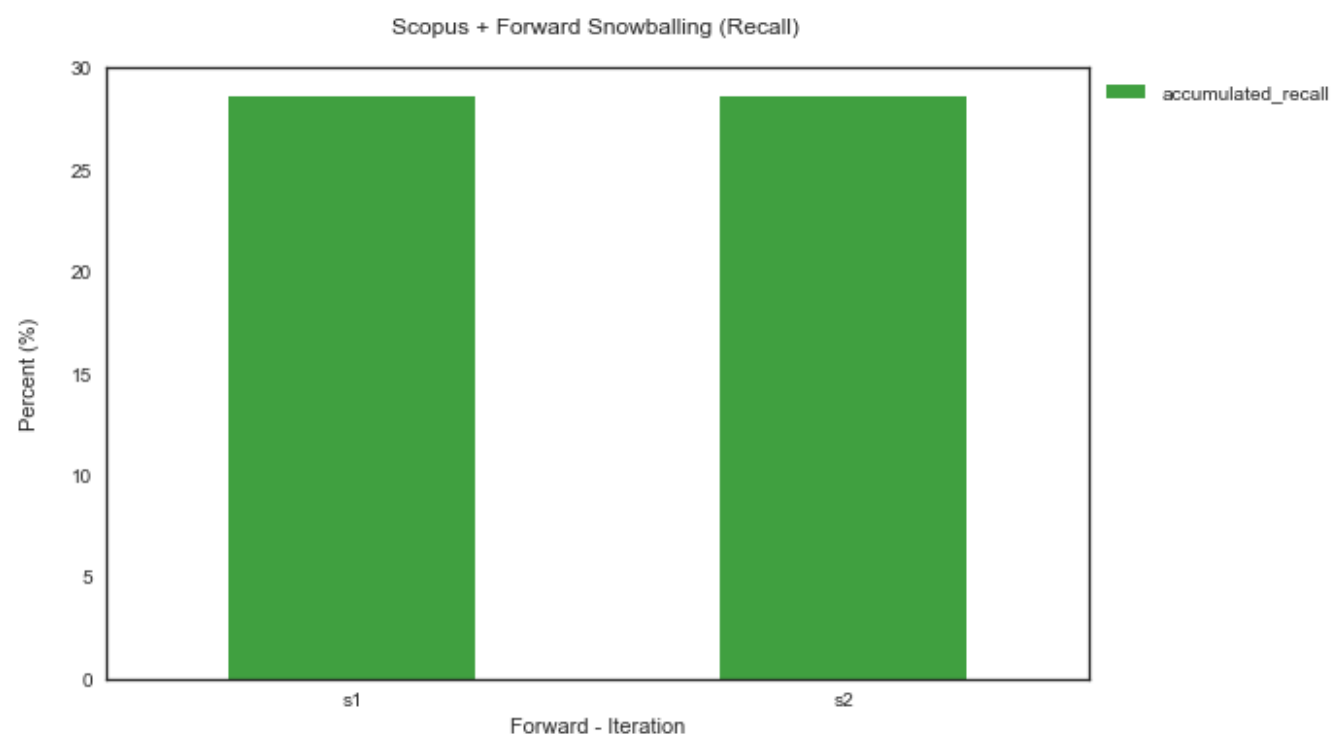
Bar Chart Strategy 5



Bar Chart Strategy 5

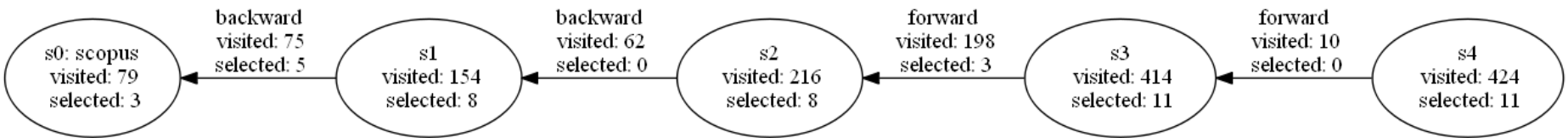


Bar Chart Strategy 5



▼ 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 = 14

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

Out[50]:

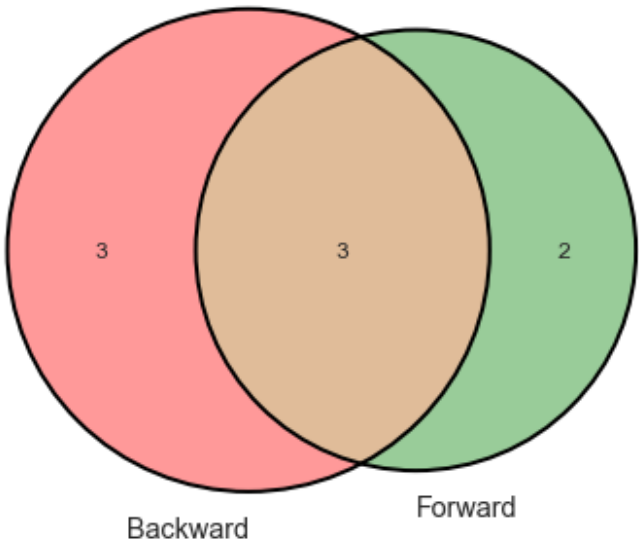
| Scopus=s0 /State | Delta Visited | Delta Selected | Delta Precision% | Delta Recall% | Delta F_Measure% | Operation | Accumulated Visited | Accumulated Selected | Accumulated Precision% | Accumulated Recall% | Accumulated F_Measure% |
|------------------|---------------|----------------|------------------|---------------|------------------|-----------|---------------------|----------------------|------------------------|---------------------|------------------------|
| s0: scopus | 79 | 3 | 3.80 | 21.43 | 6.45 | - | 79 | 3 | 3.80 | 21.43 | 6.45 |
| s1 | 75 | 5 | 6.67 | 35.71 | 11.24 | backward | 154 | 8 | 5.19 | 57.14 | 9.52 |
| s2 | 62 | 0 | 0.00 | 0.00 | 0.00 | backward | 216 | 8 | 3.70 | 57.14 | 6.96 |
| s3 | 198 | 3 | 1.52 | 21.43 | 2.83 | forward | 414 | 11 | 2.66 | 78.57 | 5.14 |
| s4 | 10 | 0 | 0.00 | 0.00 | 0.00 | forward | 424 | 11 | 2.59 | 78.57 | 5.02 |

Report Strategy 6 - Measures of Strategy (accumulated)

Out[51]:

| Strategy | Visited | Selected | Precision % | Recall % | F-Measure % |
|------------------|---------|----------|-------------|----------|-------------|
| Scopus + BS + FS | 424 | 11 | 2.59 | 78.57 | 5.02 |

Diagrama de Venn - Strategy 6

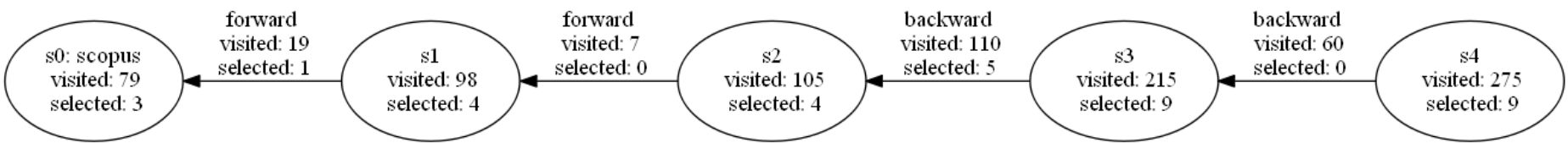


Report Diagrama de Venn

Backward ['grambow2012a', 'soydan2006a', 'rungratri2008a', 'lee2008a', 'lee2009a', 'liao2005a']
Forward ['grambow2013a', 'rungratri2008a', 'grambow2012a', 'soydan2006a', 'soydan2012a']

▼ 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 = 14

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

Out[56]:

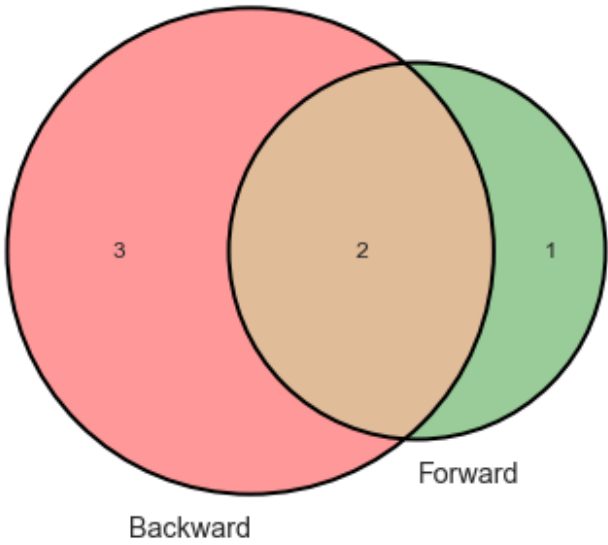
| Scopus=s0 /State | Delta Visited | Delta Selected | Delta Precision% | Delta Recall% | Delta F_Measure% | Operation | Accumulated Visited | Accumulated Selected | Accumulated Precision% | Accumulated Recall% | Accumulated F_Measure% |
|------------------|---------------|----------------|------------------|---------------|------------------|-----------|---------------------|----------------------|------------------------|---------------------|------------------------|
| s0: scopus | 79 | 3 | 3.80 | 21.43 | 6.45 | - | 79 | 3 | 3.80 | 21.43 | 6.45 |
| s1 | 19 | 1 | 5.26 | 7.14 | 6.06 | forward | 98 | 4 | 4.08 | 28.57 | 7.14 |
| s2 | 7 | 0 | 0.00 | 0.00 | 0.00 | forward | 105 | 4 | 3.81 | 28.57 | 6.72 |
| s3 | 110 | 5 | 4.55 | 35.71 | 8.06 | backward | 215 | 9 | 4.19 | 64.29 | 7.86 |
| s4 | 60 | 0 | 0.00 | 0.00 | 0.00 | backward | 275 | 9 | 3.27 | 64.29 | 6.23 |

Report Strategy 7 - Measures of Strategy (accumulated)

Out[57]:

| Strategy | Visited | Selected | Precision % | Recall % | F-Measure % |
|------------------|---------|----------|-------------|----------|-------------|
| Scopus + FS + BS | 275 | 9 | 3.27 | 64.29 | 6.23 |

Diagrama de Venn - Strategy 7



Report Diagrama de Venn

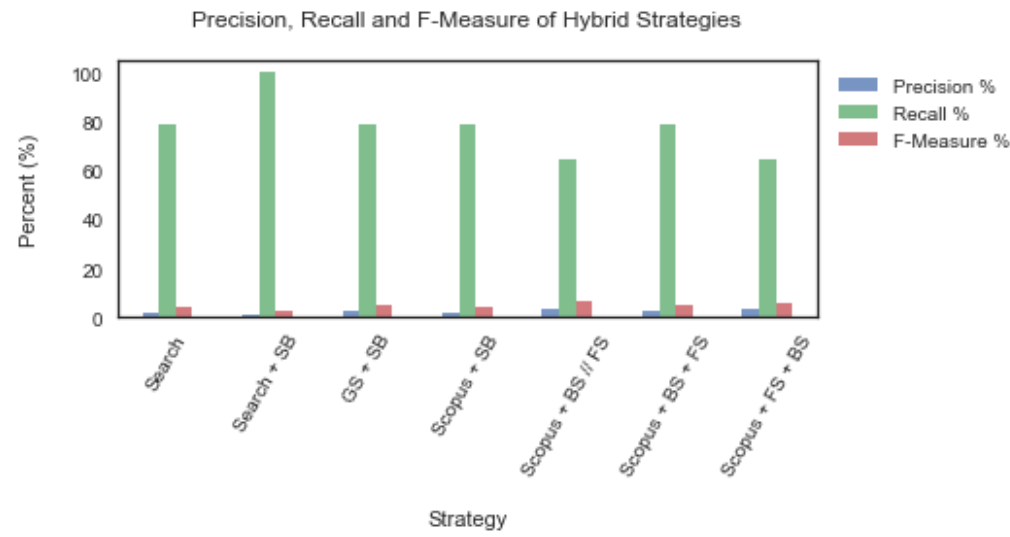
Backward ['soydan2006a', 'rungratri2008a', 'lee2008a', 'liao2005a', 'lee2009a']
Forward ['soydan2012a', 'rungratri2008a', 'soydan2006a']

▼ 3 Analyses

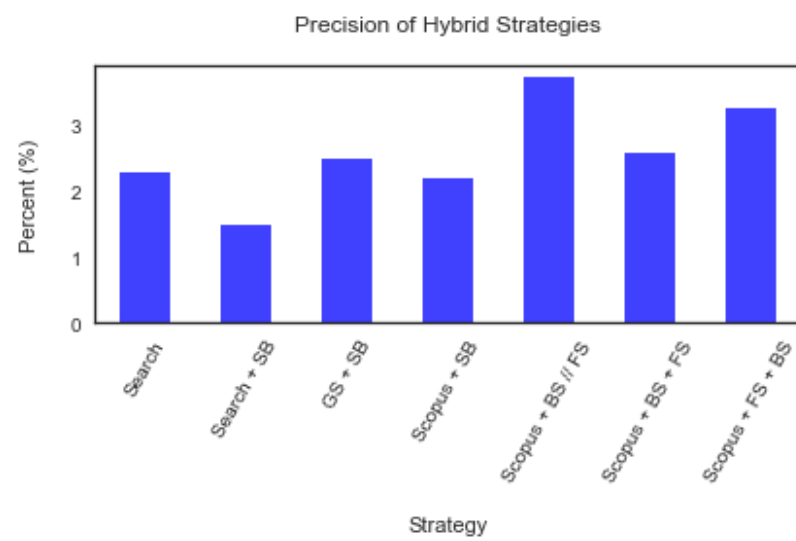
Out[61]:

| Strategy | Visited | Selected | Precision % | Recall % | F-Measure % |
|-------------------|---------|----------|-------------|----------|-------------|
| Search | 480 | 11 | 2.29 | 78.57 | 4.45 |
| Search + SB | 932 | 14 | 1.50 | 100.00 | 2.96 |
| GS + SB | 442 | 11 | 2.49 | 78.57 | 4.82 |
| Scopus + SB | 502 | 11 | 2.19 | 78.57 | 4.26 |
| Scopus + BS // FS | 242 | 9 | 3.72 | 64.29 | 7.03 |
| Scopus + BS + FS | 424 | 11 | 2.59 | 78.57 | 5.02 |
| Scopus + FS + BS | 275 | 9 | 3.27 | 64.29 | 6.23 |

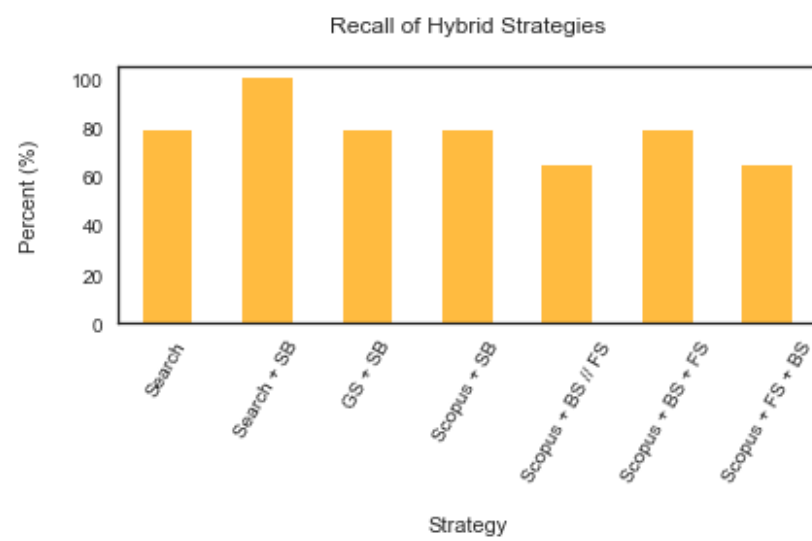
Bar Chart Strategies



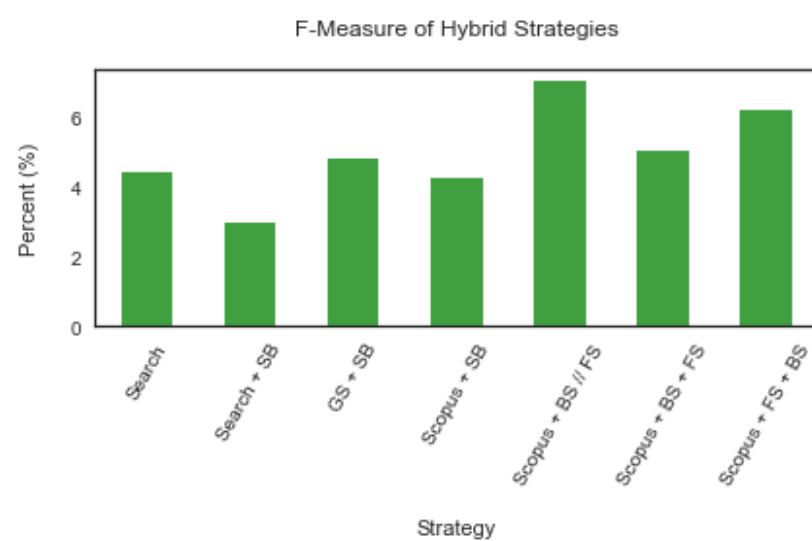
Bar Chart Strategies



Bar Chart Strategies



Bar Chart Strategies



▼ 4 Research Questions

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

Out[67]:

| Digital Library | Visited | Selected | Precision % | Recall % | F-Measure % |
|-----------------|---------|----------|-------------|----------|-------------|
| acm | 0 | 0 | nan | 0.00 | 0.00 |
| googlescholar | 466 | 11 | 2.36 | 78.57 | 4.58 |
| ieee | 0 | 0 | nan | 0.00 | 0.00 |
| sciencedirect | 21 | 0 | 0.00 | 0.00 | 0.00 |
| scopus | 79 | 3 | 3.80 | 21.43 | 6.45 |
| springer | 65 | 1 | 1.54 | 7.14 | 2.53 |
| webofscience | 0 | 0 | nan | 0.00 | 0.00 |
| wiley | 15 | 0 | 0.00 | 0.00 | 0.00 |

RQ2) Which of the hybrid strategies is more efficient?

Out[68]:

| Strategy | Precision % | Recall % | F-Measure % |
|-------------------|-------------|----------|-------------|
| Search + SB | 1.50 | 100.00 | 2.96 |
| GS + SB | 2.49 | 78.57 | 4.82 |
| Scopus + SB | 2.19 | 78.57 | 4.26 |
| Scopus + BS // FS | 3.72 | 64.29 | 7.03 |
| Scopus + BS + FS | 2.59 | 78.57 | 5.02 |
| Scopus + FS + BS | 3.27 | 64.29 | 6.23 |

RQ3) Which of the strategies is more efficient?

Out[69]:

| Strategy | Precision % | Recall % | F-Measure % |
|-------------------|-------------|----------|-------------|
| Search | 2.29 | 78.57 | 4.45 |
| Search + SB | 1.50 | 100.00 | 2.96 |
| GS + SB | 2.49 | 78.57 | 4.82 |
| Scopus + SB | 2.19 | 78.57 | 4.26 |
| Scopus + BS // FS | 3.72 | 64.29 | 7.03 |
| Scopus + BS + FS | 2.59 | 78.57 | 5.02 |
| Scopus + FS + BS | 3.27 | 64.29 | 6.23 |

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

Out[70]:

| Digital Library | Recall % |
|-----------------|----------|
| acm | 0.00 |
| googlescholar | 100.00 |
| ieee | 21.43 |
| sciencedirect | 7.14 |
| scopus | 50.00 |
| springer | 21.43 |
| webofscience | 35.71 |
| wiley | 0.00 |