1 Investigating Hybrid Strategies for Systematic Literature Review
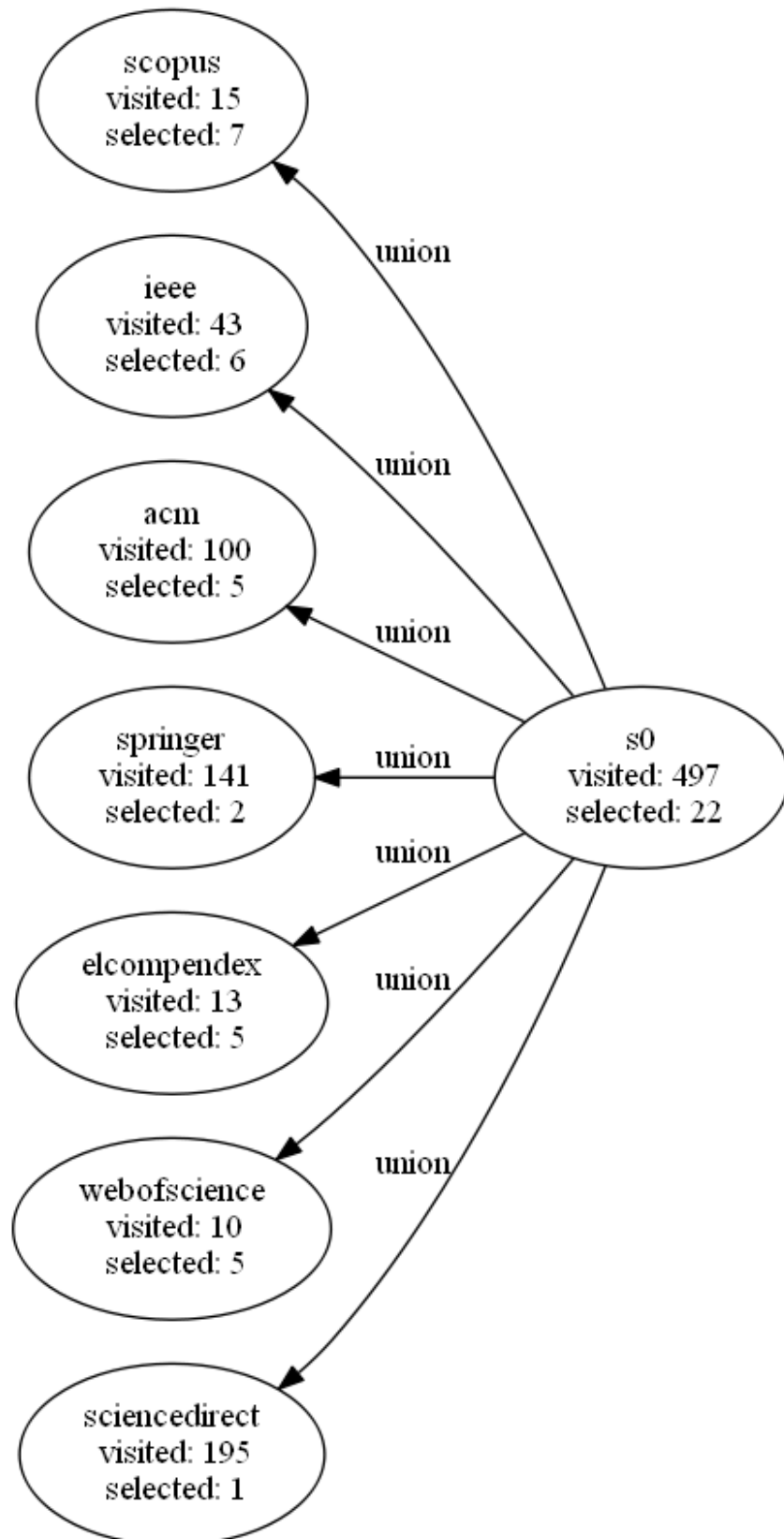
Experiment 1

2 Strategies

2.1 Strategy 1 - Database Search

Total de estudos primários reproduzidos = 22 do Seed Set + 29 encontrados durante o Snowballing = 51
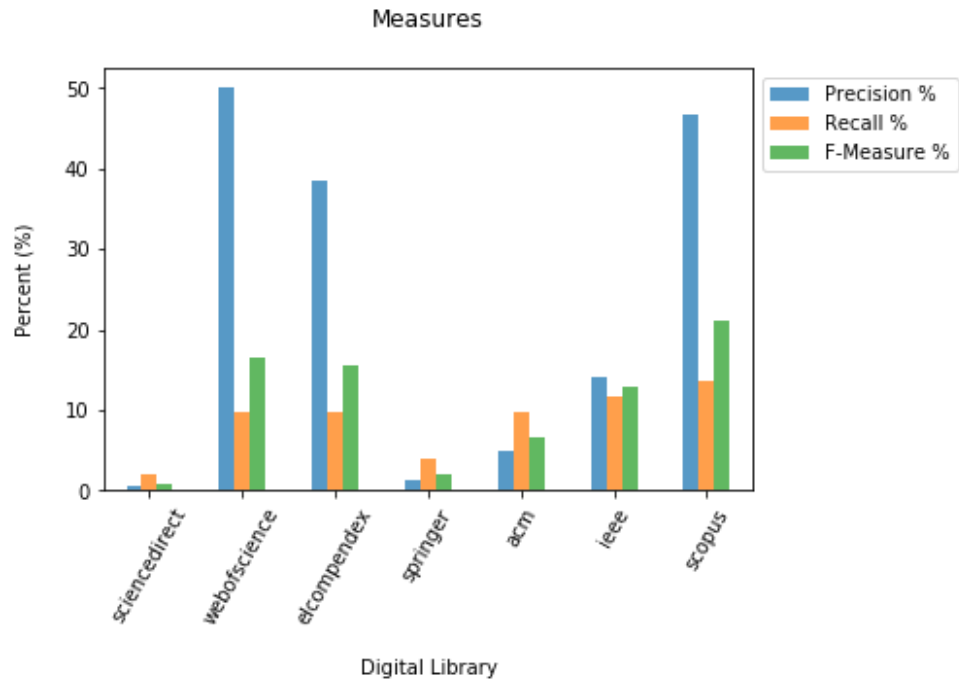
Provenance Graph - Strategy 1 - Database Search

Report Strategy 1 - Database Search

Digital Library	Precision %	Recall %	F-Measure %
sciencedirect	0.5	2	0.8
webofscience	50	9.8	16.4
elcompendex	38.5	9.8	15.6
springer	1.4	3.9	2.1
acm	5	9.8	6.6
ieee	14	11.8	12.8
scopus	46.7	13.7	21.2

Bar Chart Strategy 1

Measures

Report Strategy 1 Measures (Accumulated)

Precision %	Recall %	F-Measure %	Visited	Selected
4.4	43.1	8	497	22

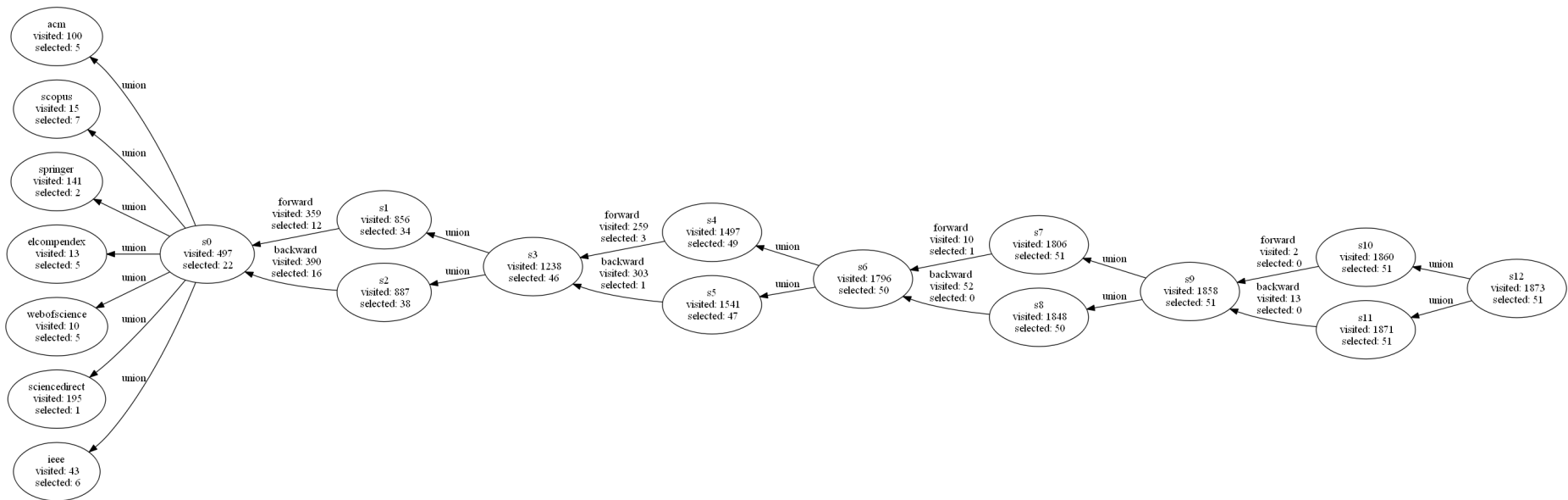**Bar Chart Publications Type - Strategy 1**

Publication Type of Selected

Publication Year of Selected
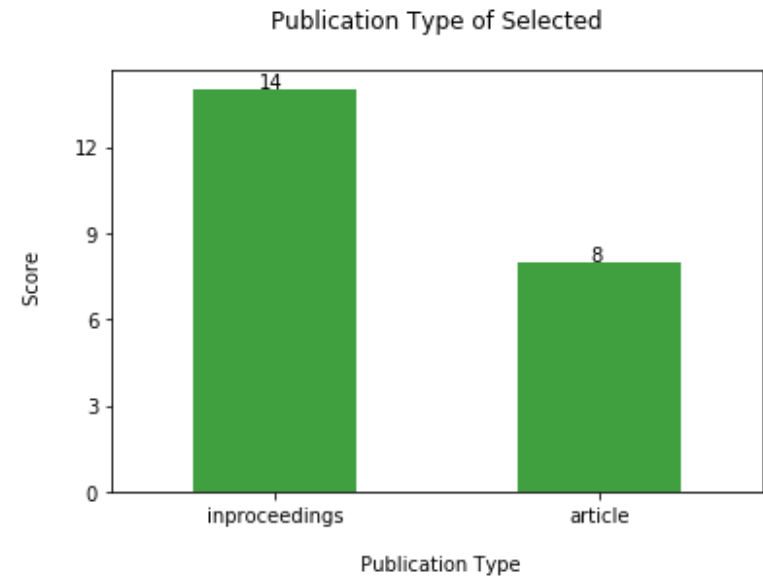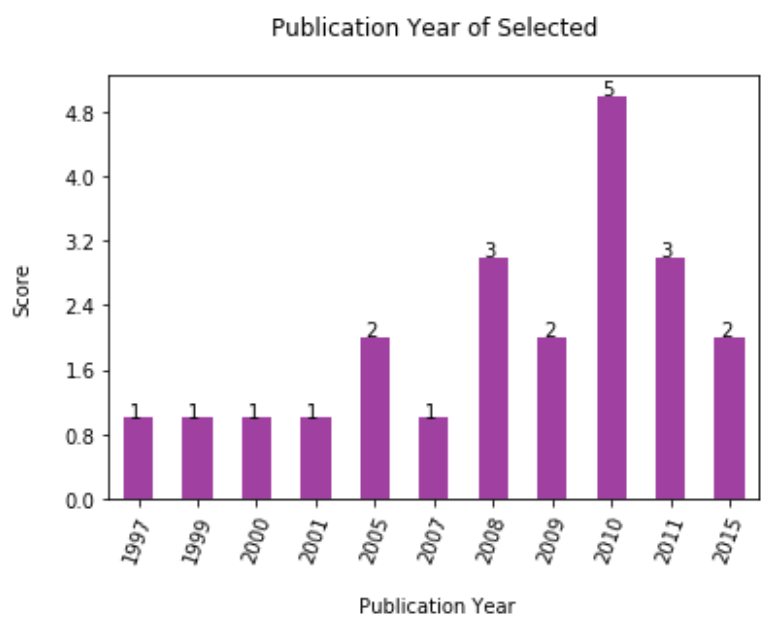
2.2 Strategy 2 - Database Search + Snowballing (Guideline Wohlin)

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

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

Out[16]:

Digital Library/State	Precision%	Recall%	FMeasure%	Operation	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%	Delta Visited	Visited	Delta Related	Related
ieee	14	11.8	12.8	-	14	11.8	12.8	43	43	6	6
sciencedirect	0.5	2	0.8	-	0.5	2	0.8	195	195	1	1
webofscience	50	9.8	16.4	-	50	9.8	16.4	10	10	5	5
elcompendex	38.5	9.8	15.6	-	38.5	9.8	15.6	13	13	5	5
springer	1.4	3.9	2.1	-	1.4	3.9	2.1	141	141	2	2
scopus	46.7	13.7	21.2	-	46.7	13.7	21.2	15	15	7	7
acm	5	9.8	6.6	-	5	9.8	6.6	100	100	5	5
s0	nan	0	0	union	4.4	43.1	8	0	497	0	22
s2	4.1	31.4	7.3	backward	4.3	74.5	8.1	390	887	16	38
s1	3.3	23.5	5.9	forward	4	66.7	7.5	359	856	12	34
s3	nan	0	0	union	3.7	90.2	7.1	0	1238	0	46
s5	0.3	2	0.6	backward	3	92.2	5.9	303	1541	1	47
s4	1.2	5.9	1.9	forward	3.3	96.1	6.3	259	1497	3	49
s6	nan	0	0	union	2.8	98	5.4	0	1796	0	50
s8	0	0	0	backward	2.7	98	5.3	52	1848	0	50
s7	10	2	3.3	forward	2.8	100	5.5	10	1806	1	51
s9	nan	0	0	union	2.7	100	5.3	0	1858	0	51
s11	0	0	0	backward	2.7	100	5.3	13	1871	0	51
s10	0	0	0	forward	2.7	100	5.3	2	1860	0	51
s12	nan	0	0	union	2.7	100	5.3	0	1873	0	51

```
Report Strategy 2 - Measures (Accumulated)
```

| Precision %	Recall %	F-Measure %	Visited	Selected
2.7	100	5.3	1873	51

Diagrama de Venn - Strategy 2



```
Report Diagrama de Venn
```
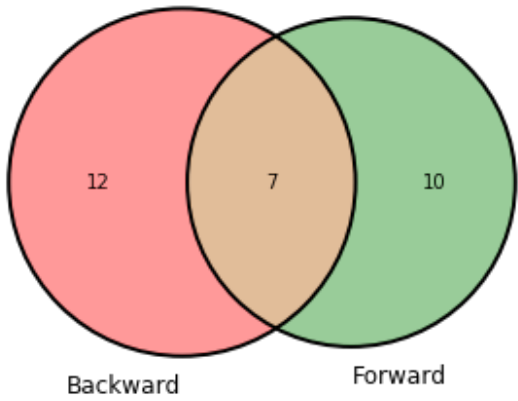
Backward ['liu2005a', 'münch2013a', 'mccoy1998a', 'martins2007b', 'kowalczyk2010a', 'debou2000a', 'basili2007c', 'münch2013c', 'liu2006a', 'mcloughlin2010a', 'mandić2010d', 'basili2007a', 'trendowicz2011a', 'basili2007b', 'lepmets2014a', 'armbrust2010a', 'murugappan2003a', 'messnarz1999a', 'hinley1995a']
Forward ['trienekens2004a', 'armbrust2010b', 'münch2013a', 'kowalczyk2010a', 'stallinger2013a', 'münch2013c', 'mandić2010d', 'mcloughlin2010a', 'sun2010a', 'unterkalmsteiner2014a', 'mcloughlin2010b', 'karlström2002a', 'trendowicz2011a', 'armbrust2010a', 'birkhölzer2011a', 'cocozza2014a', 'trendowicz2014a']
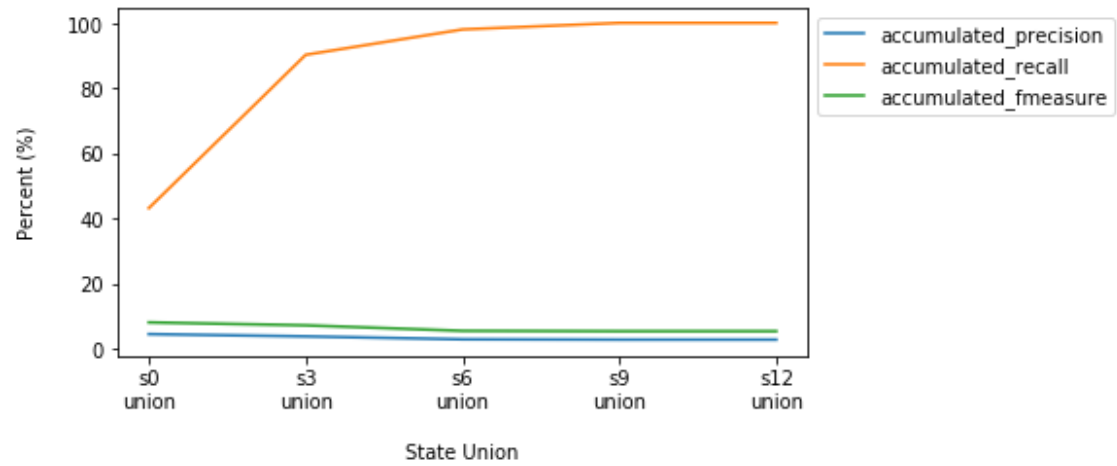
2.3 Strategy 3 - Informal Search (Google Scholar) + Snowballing (Guideline Wohlin)

```
Provenance Graph - Strategy 3 - Informal Search (Google Scholar) + Snowballing (Guideline Wohlin)
```
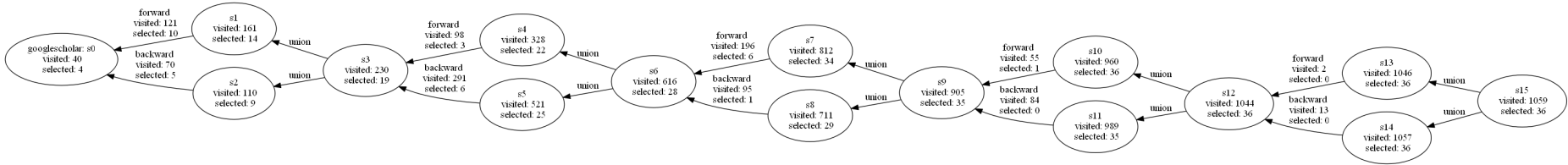
```
Report Strategy 3 - Informal Search (Google Scholar) + Snowballing (Guideline Wohlin)
```

Out[23]:

Scholar=s0 /State	Precision%	Recall%	FMeasure%	Operation	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%	Delta Visited	Visited	Delta Related	Related
googlescholar: s0	10	7.8	8.8	-	10	7.8	8.8	40	40	4	4
s2	7.1	9.8	8.3	backward	8.2	17.6	11.2	70	110	5	9
s1	8.3	19.6	11.6	forward	8.7	27.5	13.2	121	161	10	14
s3	nan	0	0	union	8.3	37.3	13.5	0	230	0	19
s5	2.1	11.8	3.5	backward	4.8	49	8.7	291	521	6	25
s4	3.1	5.9	4	forward	6.7	43.1	11.6	98	328	3	22
s6	nan	0	0	union	4.5	54.9	8.4	0	616	0	28
s8	1.1	2	1.4	backward	4.1	56.9	7.6	95	711	1	29
s7	3.1	11.8	4.9	forward	4.2	66.7	7.9	196	812	6	34
s9	nan	0	0	union	3.9	68.6	7.3	0	905	0	35
s11	0	0	0	backward	3.5	68.6	6.7	84	989	0	35
s10	1.8	2	1.9	forward	3.7	70.6	7.1	55	960	1	36
s12	nan	0	0	union	3.4	70.6	6.6	0	1044	0	36
s14	0	0	0	backward	3.4	70.6	6.5	13	1057	0	36
s13	0	0	0	forward	3.4	70.6	6.6	2	1046	0	36
s15	nan	0	0	union	3.4	70.6	6.5	0	1059	0	36

```
Report Strategy 3 - Measures (Accumulated)
```

Out[24]:

Precision %	Recall %	F-Measure %	Visited	Selected
3.4	70.6	6.5	1059	36

Diagrama de Venn - Strategy 3

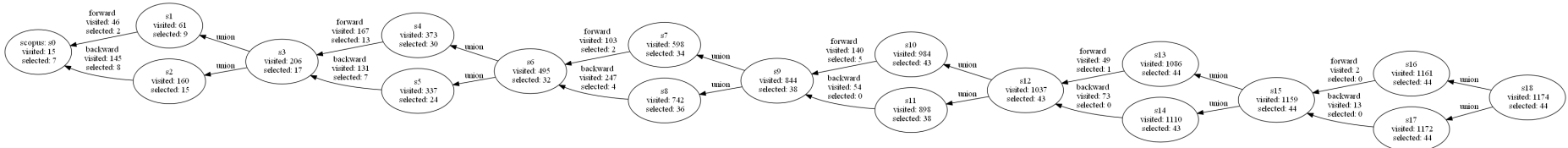Backward   10   9   13   Forward

```
Report Diagrama de Venn
```
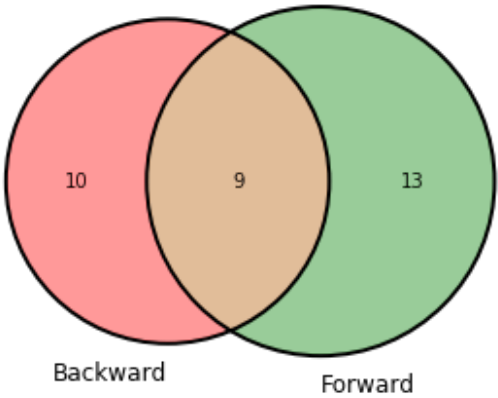
```
Backward ['trienekens2009a', 'liu2005a', 'münch2013a', 'mccoy1998a', 'kowalczyk2010a', 'debou2000a', 'basili2007c', 'münch2
013c', 'liu2006a', 'mcloughlin2010a', 'mandić2010d', 'becker2008b', 'basili2007a', 'trendowicz2011a', 'basili2007b', 'armbr
ust2010a', 'murugappan2003a', 'messnarz1999a', 'mandić2010a']
Forward ['mandić2010d', 'mcloughlin2010a', 'wang2005a', 'unterkalmsteiner2014a', 'trendowicz2011a', 'trienekens2004a', 'mün
ch2013a', 'münch2013c', 'sun2010a', 'mcloughlin2010b', 'guzmán2010a', 'petersen2015a', 'becker2008b', 'mandić2010a', 'trend
owicz2014a', 'becker2008a', 'armbrust2010b', 'kowalczyk2010a', 'armbrust2010a', 'birkhölzer2011a', 'cocozza2014a', 'trienek
ens2005a']
```

2.4 Strategy 4 - Scopus + Snowballing (Guideline Wohlin)

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

Out[29]:

Scopus=s0 /State	Precision%	Recall%	FMeasure%	Operation	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%	Delta Visited	Visited	Delta Related	Related
scopus: s0	46.7	13.7	21.2	-	46.7	13.7	21.2	15	15	7	7
s2	5.5	15.7	8.2	backward	9.4	29.4	14.2	145	160	8	15
s1	4.3	3.9	4.1	forward	14.8	17.6	16.1	46	61	2	9
s3	nan	0	0	union	8.3	33.3	13.2	0	206	0	17
s5	5.3	13.7	7.7	backward	7.1	47.1	12.4	131	337	7	24
s4	7.8	25.5	11.9	forward	8	58.8	14.2	167	373	13	30
s6	nan	0	0	union	6.5	62.7	11.7	0	495	0	32
s8	1.6	7.8	2.7	backward	4.9	70.6	9.1	247	742	4	36
s7	1.9	3.9	2.6	forward	5.7	66.7	10.5	103	598	2	34
s9	nan	0	0	union	4.5	74.5	8.5	0	844	0	38
s11	0	0	0	backward	4.2	74.5	8	54	898	0	38
s10	3.6	9.8	5.2	forward	4.4	84.3	8.3	140	984	5	43
s12	nan	0	0	union	4.1	84.3	7.9	0	1037	0	43
s14	0	0	0	backward	3.9	84.3	7.4	73	1110	0	43
s13	2	2	2	forward	4.1	86.3	7.7	49	1086	1	44
s15	nan	0	0	union	3.8	86.3	7.3	0	1159	0	44
s17	0	0	0	backward	3.8	86.3	7.2	13	1172	0	44
s16	0	0	0	forward	3.8	86.3	7.3	2	1161	0	44
s18	nan	0	0	union	3.7	86.3	7.2	0	1174	0	44

Out[30]:

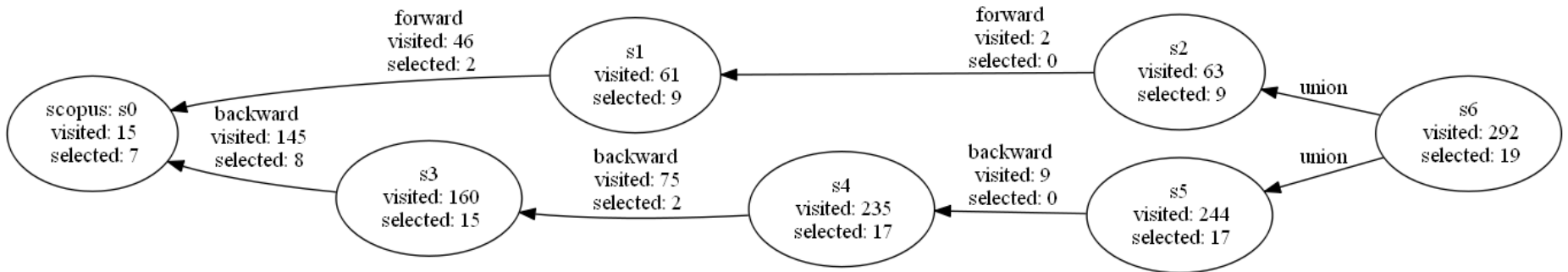| Precision %	Recall %	F-Measure %	Visited	Selected
3.7	86.3	7.2	1174	44

Diagrama de Venn - Strategy 4
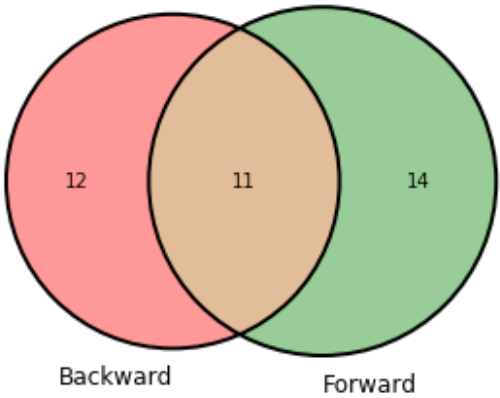
Report Diagrama de Venn

Backward ['liu2005a', 'basili2007c', 'mcloughlin2010a', 'mandić2010d', 'basili2010a', 'trendowicz2011a', 'trienekens2009a', 'münch2013a', 'münch2013c', 'liu2006a', 'murugappan2003a', 'mandić2010b', 'becker2008b', 'basili2007a', 'basili2007b', 'mandić2010a', 'messnarz1999a', 'mccoy1998a', 'martins2007b', 'kowalczyk2010a', 'debou2000a', 'armbrust2010a', 'hinley1995a']
Forward ['mandić2010d', 'mcloughlin2010a', 'wang2005a', 'unterkalmsteiner2014a', 'trendowicz2011a', 'basili2010a', 'barreto2010a', 'trienekens2004a', 'münch2013a', 'münch2013c', 'sun2010a', 'mcloughlin2010b', 'martins2008a', 'mandić2010b', 'petersen2015a', 'becker2008b', 'mandić2010a', 'trendowicz2014a', 'armbrust2010b', 'kowalczyk2010a', 'stallinger2013a', 'armbrust2010a', 'birkhölzer2011a', 'cocozza2014a', 'trienekens2005a']

▼ 2.5 Strategy 5 - Scopus + Snowballing (Short Paper)

Provenance Graph - Strategy 5 - Scopus + Snowballing (Short Paper)

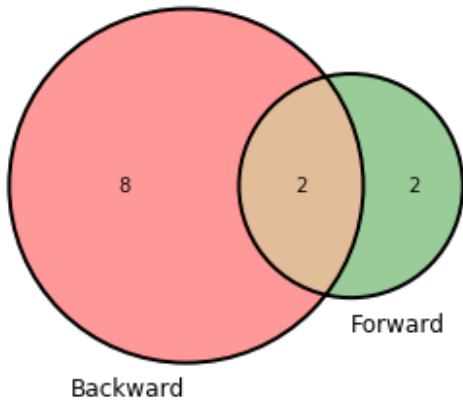**Report Strategy 5 - Scopus + Snowballing (Short Paper)**

Out[35]:

Scopus=s0 /State	Precision%	Recall%	FMeasure%	Operation	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%	Delta Visited	Visited	Delta Related	Related
s3	5.5	15.7	8.2	backward	9.4	29.4	14.2	145	160	8	15
scopus: s0	46.7	13.7	21.2	-	46.7	13.7	21.2	15	15	7	7
s4	2.7	3.9	3.2	backward	7.2	33.3	11.9	75	235	2	17
s1	4.3	3.9	4.1	forward	14.8	17.6	16.1	46	61	2	9
s5	0	0	0	backward	7	33.3	11.5	9	244	0	17
s2	0	0	0	forward	14.3	17.6	15.8	2	63	0	9
s6	nan	0	0	union	6.5	37.3	11.1	0	292	0	19

Report Strategy 5 - Measures (Accumulated)

Out[36]:

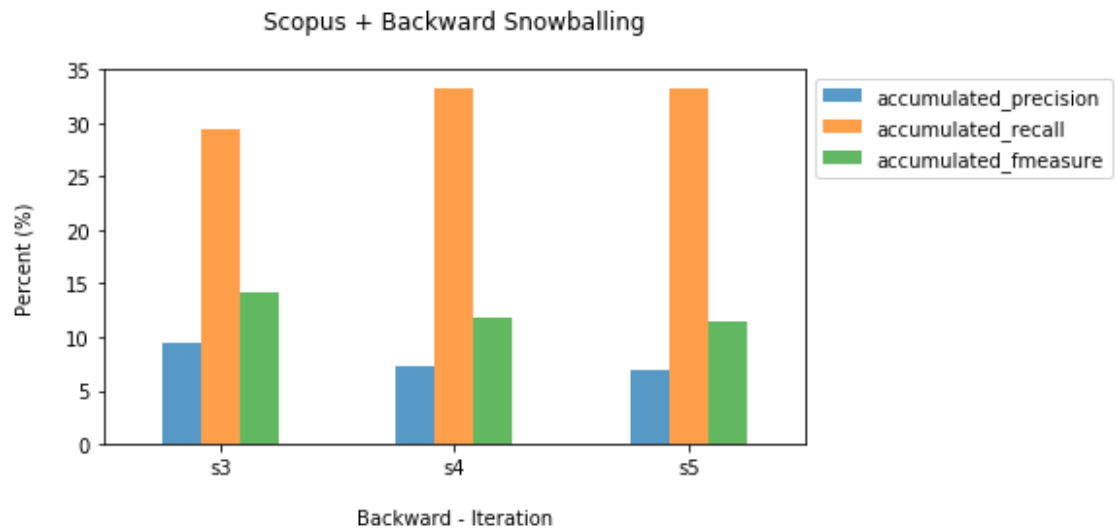| Precision %	Recall %	F-Measure %	Visited	Selected
6.5	37.3	11.1	292	19

Diagrama de Venn - Strategy 5

Report Diagrama de Venn

Backward ['trienekens2009a', 'liu2005a', 'mccoy1998a', 'martins2007b', 'kowalczyk2010a', 'basili2007c', 'basili2007a', 'basili2010a', 'basili2007b', 'hinley1995a']
Forward ['petersen2015a', 'basili2010a', 'kowalczyk2010a', 'stallinger2013a']
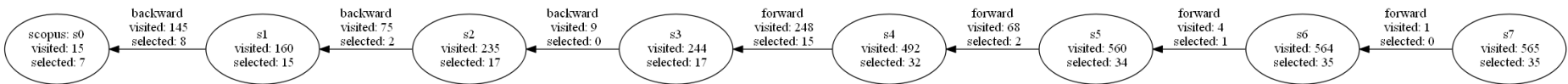
Bar Chart Strategy 5

Bar Chart Strategy 5
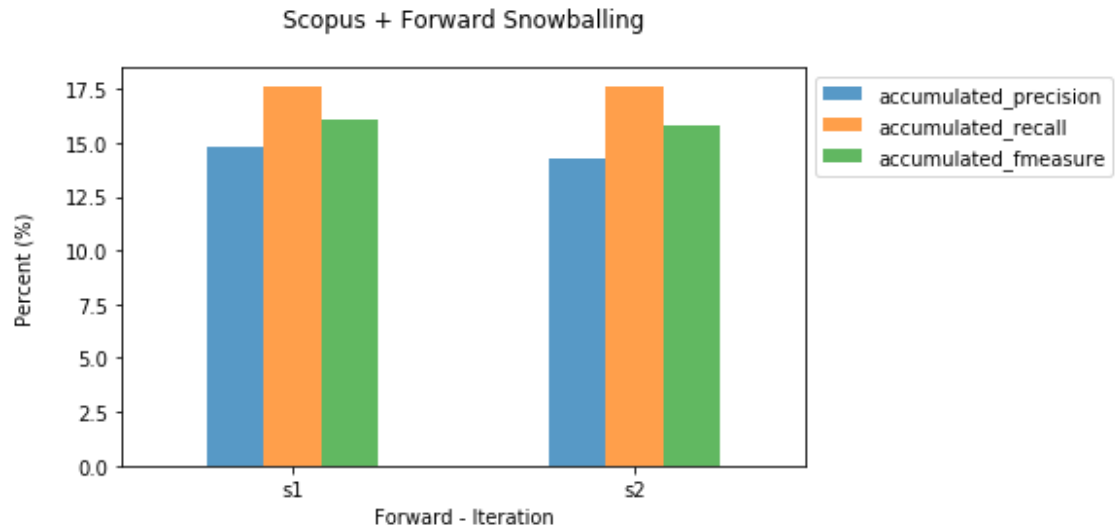
2.6 Strategy 6 - Scopus + Snowballing (JF BF)

Provenance Graph - Strategy 6 - Scopus + Snowballing (JF BF)

Report Strategy 6 - Scopus + Snowballing (JF BF)

Out[43]:

Scopus=s0 /State	Precision%	Recall%	FMeasure%	Operation	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%	Delta Visited	Visited	Delta Related	Related
scopus: s0	46.7	13.7	21.2	-	46.7	13.7	21.2	15	15	7	7
s1	5.5	15.7	8.2	backward	9.4	29.4	14.2	145	160	8	15
s2	2.7	3.9	3.2	backward	7.2	33.3	11.9	75	235	2	17
s3	0	0	0	backward	7	33.3	11.5	9	244	0	17
s4	6	29.4	10	forward	6.5	62.7	11.8	248	492	15	32
s5	2.9	3.9	3.4	forward	6.1	66.7	11.1	68	560	2	34
s6	25	2	3.6	forward	6.2	68.6	11.4	4	564	1	35
s7	0	0	0	forward	6.2	68.6	11.4	1	565	0	35

Report Strategy 6 - Measures (Accumulated)

Out[44]:

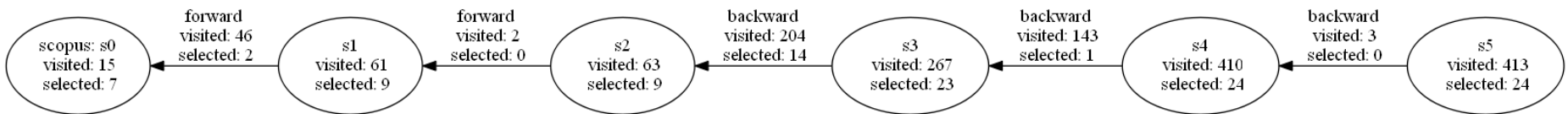| Precision %	Recall %	F-Measure %	Visited	Selected
6.2	68.6	11.4	565	35

Diagrama de Venn - Strategy 6
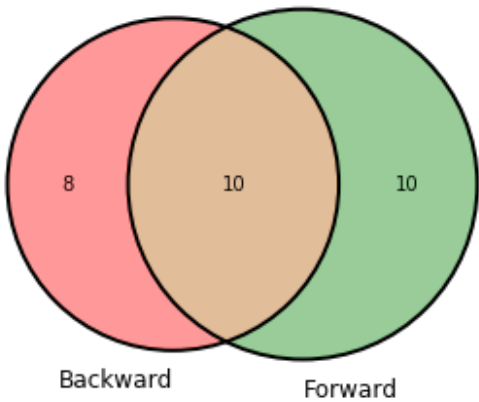
Report Diagrama de Venn

Backward ['trienekens2009a', 'liu2005a', 'münch2013a', 'mccoy1998a', 'martins2007b', 'kowalczyk2010a', 'mandić2010b', 'basili2007c', 'münch2013c', 'mandić2010d', 'becker2008b', 'basili2007a', 'basili2010a', 'basili2007b', 'trendowicz2011a', 'armbrust2010a', 'mandić2010a', 'hinley1995a']
Forward ['mandić2010d', 'unterkalmsteiner2014a', 'trendowicz2011a', 'basili2010a', 'barreto2010a', 'münch2013a', 'münch2013c', 'sun2010a', 'martins2008a', 'mandić2010b', 'petersen2015a', 'becker2008b', 'mandić2010a', 'trendowicz2014a', 'armbrust2010b', 'kowalczyk2010a', 'stallinger2013a', 'armbrust2010a', 'birkhölzer2011a', 'cocozza2014a']

2.7 Strategy 7 - Scopus + Snowballing (JF FB)

Provenance Graph - Strategy 7 - Scopus + Snowballing (JF FB)

**Report Strategy 7 - Scopus + Snowballing (JF FB)**

Out[49]:

Scopus=s0 /State	Precision%	Recall%	FMeasure%	Operation	Accumulated Precision%	Accumulated Recall%	Accumulated F_Measure%	Delta Visited	Visited	Delta Related	Related
scopus: s0	46.7	13.7	21.2	-	46.7	13.7	21.2	15	15	7	7
s1	4.3	3.9	4.1	forward	14.8	17.6	16.1	46	61	2	9
s2	0	0	0	forward	14.3	17.6	15.8	2	63	0	9
s3	6.9	27.5	11	backward	8.6	45.1	14.5	204	267	14	23
s4	0.7	2	1	backward	5.9	47.1	10.4	143	410	1	24
s5	0	0	0	backward	5.8	47.1	10.3	3	413	0	24

Report Strategy 7 - Measures (Accumulated)

Out[50]:

Precision %	Recall %	F-Measure %	Visited	Selected
5.8	47.1	10.3	413	24


Diagrama de Venn - Strategy 7
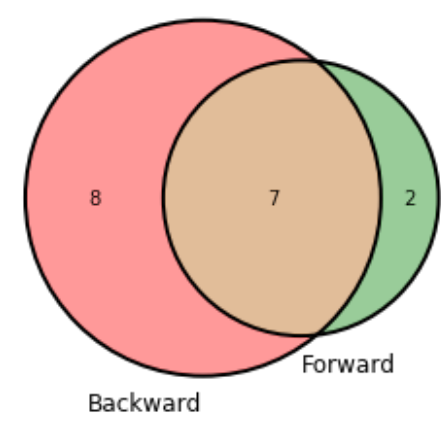
Report Diagrama de Venn

Backward ['trienekens2009a', 'liu2005a', 'münch2013a', 'mccoy1998a', 'martins2007b', 'kowalczyk2010a', 'mandić2010b', 'basili2007c', 'münch2013c', 'mandić2010d', 'basili2007a', 'basili2010a', 'basili2007b', 'trendowicz2011a', 'hinley1995a']
Forward ['münch2013a', 'kowalczyk2010a', 'mandić2010b', 'stallinger2013a', 'münch2013c', 'mandić2010d', 'petersen2015a', 'trendowicz2011a', 'basili2010a']

▼ 3 Analyses

Out[54]:

Strategy	Precision %	Recall %	F-Measure %	Visited	Related	Final Selected
E1	4.4	43.1	8	497	22	15
E2	2.7	100	5.3	1873	51	30
E3	3.4	70.6	6.5	1059	36	22
E4	3.7	86.3	7.2	1174	44	25
E5	6.5	37.3	11.1	292	19	14
E6	6.2	68.6	11.4	565	35	20
E7	5.8	47.1	10.3	413	24	15

Bar Chart Strategies

Precision, Recall and F-Measure of Hybrid Strategies
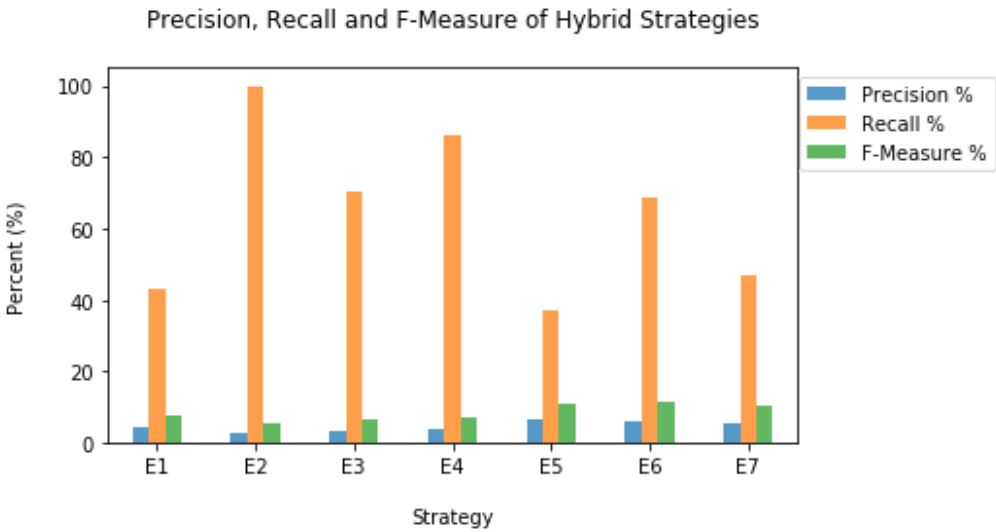
▼ 4 Research Questions

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

Out[59]:

Digital Library	Precision %	Recall %	F-Measure %
sciencedirect	0.5	2	0.8
webofscience	50	9.8	16.4
elcompendex	38.5	9.8	15.6
springer	1.4	3.9	2.1
acm	5	9.8	6.6
ieee	14	11.8	12.8
scopus	46.7	13.7	21.2

RQ2) Which of the hybrid strategies is more efficient?

Out[60]:

Strategy	Precision %	Recall %	F-Measure %
E2	2.7	100	5.3
E3	3.4	70.6	6.5
E4	3.7	86.3	7.2
E5	6.5	37.3	11.1
E6	6.2	68.6	11.4
E7	5.8	47.1	10.3

RQ3) Which of the strategies is more efficient?

Out[61]:

Strategy	Precision %	Recall %	F-Measure %
E1	4.4	43.1	8
E2	2.7	100	5.3
E3	3.4	70.6	6.5
E4	3.7	86.3	7.2
E5	6.5	37.3	11.1
E6	6.2	68.6	11.4
E7	5.8	47.1	10.3

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

Out[62]:

Digital Library	Recall %
scopus	82.4
springer	31.4
ieee	27.5
sciencedirect	3.9
acm	11.8
webofscience	52.9
elcompendex	68.6