

# Citações em rede: a visibilidade da produção dos pesquisadores da Ciência da Informação no Google Scholar.

**NATASHA MORELLO PEIXOTO:** Bolsista PIBIC/CNPq  
**SÔNIA ELISA CAREGNATO:** Orientadora

## Objetivos

Analisar a validade, o potencial e a cobertura do mecanismo de busca Google Scholar nos estudos de citação. O trabalho aqui apresentado enfoca o campo da Ciência da Informação no Brasil, no período que abrange 2004 a 2008, buscando determinar se essa produção está indexada na ferramenta e quais são os seus limites.

## Metodologia

- Seleção dos 13 pesquisadores com bolsa de produtividade em pesquisa (PQ) do CNPq, nível PQ1 ou PQSR, em 2009;
- Construção do conjunto de dados de referência, composto pelas publicações dos pesquisadores encontradas em seus respectivos Currículo Lattes;
- Coleta de dados no mecanismo Google Scholar: realizada a partir da busca do nome dos pesquisadores, utilizando várias formas de expressão do nome do autor, a fim de ampliar a cobertura;
- Comparação dos dados recuperados no Google com a lista da produção recuperada dos Lattes de cada pesquisador, a fim de validar os resultados.

## Análises dos resultados

- A forma como se empregam os operadores booleanos altera o resultado da busca por autores no Google Scholar. A que se mostrou mais apropriada foi a utilização dos operadores na escrita em inglês e em caixa alta.
- No total, foram recuperados 305 documentos, que apresentaram 580 citações. Após a checagem com os dados do Lattes, os documentos foram organizados em categorias. O resultado aparece na Figura 1.
- Os 163 documentos que compõem as classificações corretas, imprecisas, incompletas e incompleto-imprecisas, correspondem a 76 artigos, 6 orientações, 23 participações em bancas, 7 apresentações de trabalhos, 6 livros, 22 capítulos de livros, 19 anais e 3 produções diversas, conforme ilustra a Figura 2.
- Para fins de verificação da qualidade dos dados recuperados, calculou-se o índice de precisão para cada autor. Para isso, foram excluídos dados relativos à participação em bancas e orientações. O resultado aparece na Tabela 1.

## Resultados

### Classificação dos dados recuperados

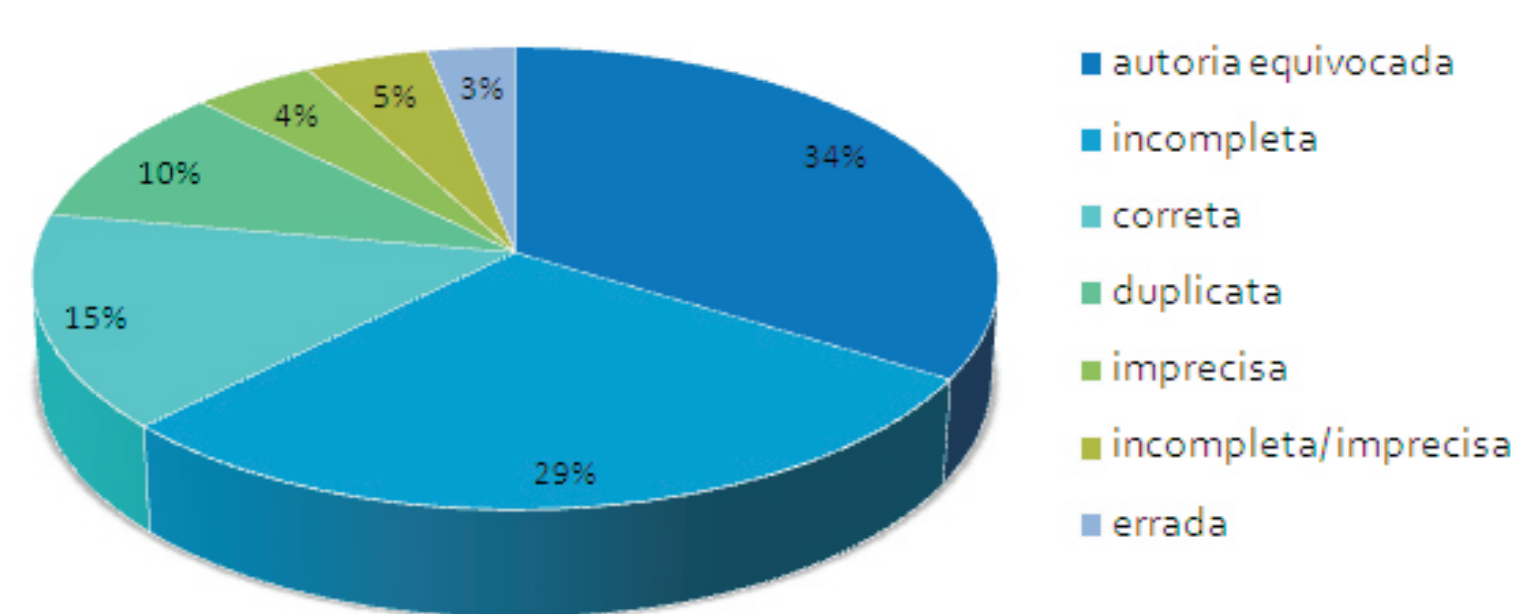


Figura 1. Classificação dos dados recuperados

### Tipologia dos documentos recuperados

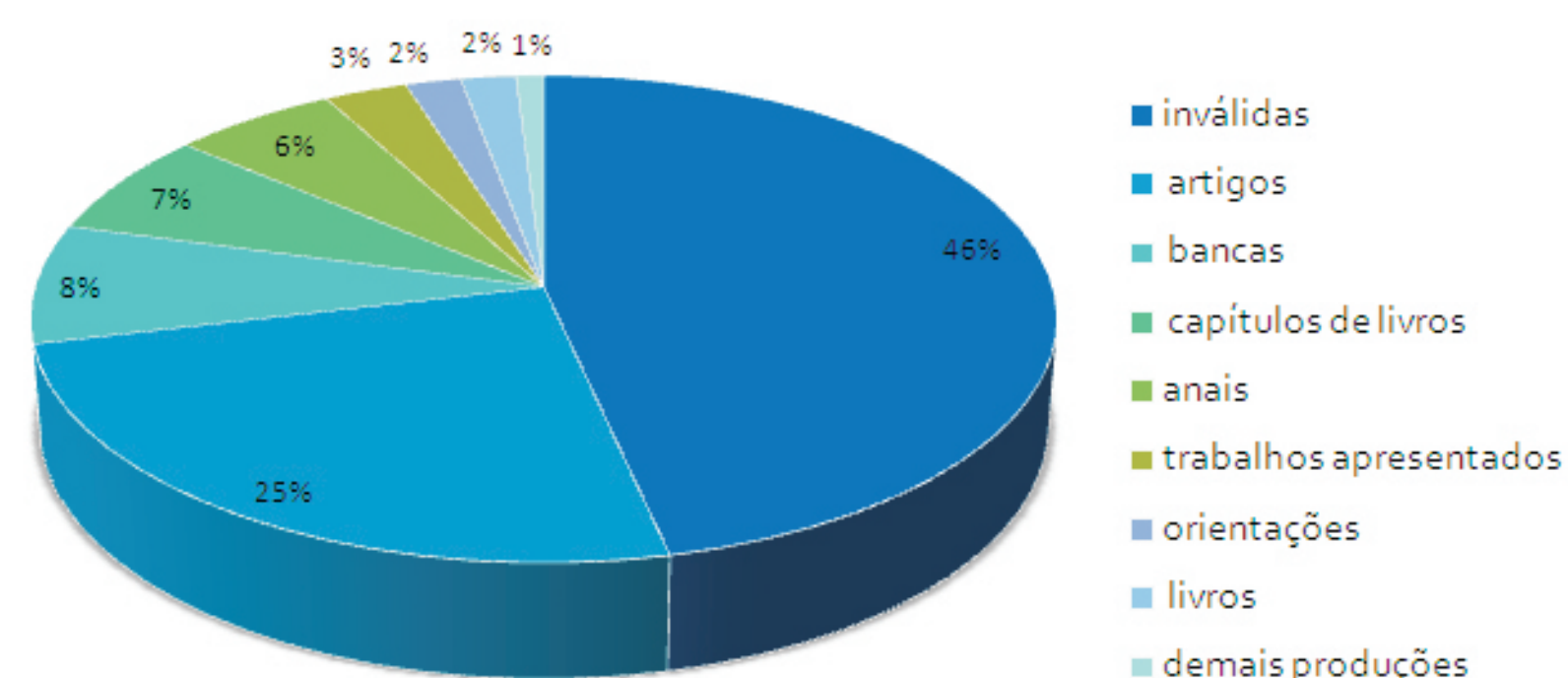


Figura 2. Tipologia dos documentos recuperados

Pesquisador	Doc. Recuperados	Docs. Úteis	%
1	20	4	20
2	9	7	78
3	21	16	77
4	24	15	62
5	64	1	1
6	7	2	30
7	16	8	50
8	34	27	80
9	23	14	60
10	8	6	75
11	17	11	60
12	20	9	45
13	42	14	33
Total	305	134	44

Tabela 1. Precisão da Recuperação

## Conclusões

- O Google Scholar mostrou ser capaz de recuperar um número grande de documentos não indexados nas bases de dados bibliográficas, tais como livros, capítulos de livros e trabalhos acadêmicos;
- Ocorrência de muitos documentos duplicados e escritos de forma incorreta, o que exige do pesquisador um tempo redobrado na limpeza de dados;
- Recomenda-se atenção na qualidade dos metadados oriundos de repositórios institucionais, e também com as análises de autoria de teses e dissertações, pois a ferramenta considera os participantes de bancas de mestrado e afins como autores, o que dificulta a análise de citações.

Apoio:



Contato: nahwhoa@yahoo.com.br