

# Systematic literature review

## A papers' analysis

Olivier Caron

Christophe Benavent

2023-05-16

## Table of contents

1	Libraries and loading data	1
2	A look at the number of marketing publications with NLP	1

## 1 Libraries and loading data

```
library(cowplot)
library(tidyverse)

#read.csv transforms ":" in "." so it was complicated to replicate code
#every column name is now lowered, the dots are replaced with _
#every underscore at the start or end of the column name is deleted

nlp_papers <- read.csv("nlp_papers_utf8.csv", fileEncoding = "UTF-8")
new_colnames <- gsub("\\\\.", "_", colnames(nlp_papers)) %>%
  trimws(., whitespace = "_") %>%
  gsub("_+", "_", .) %>%
  tolower()

colnames(nlp_papers) <- new_colnames
```

## 2 A look at the number of marketing publications with NLP

The number of papers in marketing reviews using NLP methods has increased a lot in recent years:

```

nlp_papers <- nlp_papers %>%
  mutate(year = substr(prism_coverdate,1,4))

#get rid of conference papers
nlp_papers_journal_only <- nlp_papers %>%
  filter(!grepl("conference", subtypedescription, ignore.case = TRUE))

t0 <-as.data.frame(prop.table(table(nlp_papers_journal_only$prism_publicationname)))
g01<-ggplot(t0,aes(x=reorder(Var1, Freq), y=Freq))+geom_bar(stat="identity")+
  coord_flip()+
  labs( title="Number of articles per review", y="Proportion", x= NULL)

t1<-as.data.frame(table(nlp_papers_journal_only$year))

g02<-ggplot(t1, aes(x=Var1, y=Freq, group=1))+
  geom_smooth(color="Grey70", linewidth=2)+
  geom_line(stat="identity", size=1.1) +
  labs( title="Number of publications per year", y="", x=NULL)
plotgrid <- plot_grid(g01,
                      g02,
                      labels = c('A', 'B'),
                      label_size = 10,
                      ncol=2,
                      rel_widths = c(2,1))

ggsave(filename="images/evolution_publications_nlp_marketing.png",
        width = 60,
        height = 25,
        units = "cm")

```

