Human judgements of energy trilemma discussion

Contents

Introduction

To validate the automated measure, we compared it to human judgements. 10 fluent English speakers in the UK were trained on the basics of the energy trilemma, using standard teaching resources on the topic (Glasgow Science Centre, 2021; Our Future Energy, 2022, see below). Participants were not told what the key terms were.

Each participant was asked to read 40 randomly selected articles from the corpus. For each article, they rated the extent to which each article discussed each of the three aspects of the energy trilemma, scoring each aspect independently from 0 to 10. One participant was excluded due to technical difficulties (see discussion below). 8 of the articles were identical across participants, and these were used to test the agreement between human rates.

The three measures are not independent, since discussing one aspect of the trilemma usually means not discussing the others. Therefore, the analyses below focus on sustainability, since that is the best-represented topic.

The instructions for participants are included below:

Instructions for participants

This project aims to measure how much the media discusses the 'energy trilemma': three crucial components of our global energy system. These are the topics of accessibility, security, and sustainability.

We've collected a large number of news articles about the UN conferences on climate change (e.g. COP26). We're using automatic computational linguistics methods to summarise how much discussion is occurring for each of the three areas. However, we're not sure whether the computational methods align well with human judgements. That's where you come in.

You'll read a short news article, then give your opinion about how much each part of the trilemma was discussed. We'll then use this data to check that the computational method is behaving sensibly. The first task is for you to understand what the three aspects of the trilemma are, so that you can recognise them. There's a short introduction to the topics on the next page, and a link to a short video.

Rating involves the following steps: - Open the "NewsArticles" pdf file and read one of the articles. Some of the formatting might be a little strange or the text might be cut off halfway through a sentence. Don't worry about this, just try to get a gist of what the article is about. - When you've read the article, open up the ratings excel file. You can provide your judgement about how much the article discussed each component. Make sure the article number and article ID match. - Type in a score from 0 (did not discuss) to 10 (discussed a lot) for each article. Some articles might not discuss one of the components at all, while in others there may be a balance. Looking at many articles together, we're guessing that there will be a general balance, but we could be wrong. Don't overthink things – we want to know your overall impression of each article.

Go through each of the 40 articles in the article list and rate each one. It should take about 4 hours. You're not expected to do this in one sitting.

The Energy Trilemma

We use energy to power our phones and TVs, to heat our houses, cook our food, and transport us by car, train and plane. Sources of energy include oil, gas, solar, wind etc.. The energy trilemma is about addressing three often conflicting challenges related to providing energy: ensuring energy security, providing energy accessibility, and achieving environmental sustainability.

Sustainability

Environmental Sustainability of energy systems represents the transition of a country's energy system towards mitigating and avoiding potential environmental harm and climate change impacts. The dimension focuses on productivity and efficiency of generation, transmission and distribution, decarbonisation, and air quality. Globally, we draw most of our energy from oil, coal, and natural gas. These fossil fuels account for 80% of the world's energy mix. These sources of energy have negative effects on our planet by releasing greenhouse gases into the atmosphere and are a huge contributor to the climate crisis. Sustainable energy focuses on meeting the energy demands of today without negatively impacting future generations. Hydro, solar, and wind power are all considered more sustainable sources of energy as they come from renewable sources. Other low-carbon options, such as nuclear power, may be a big part of our energy mix in the future but there are still ongoing debates about its sustainability when it comes to nuclear waste.

Security

Security refers to whether we are able to access enough energy when and where we need it. This means being able to have uninterrupted availability of energy. In the short term this could mean an energy system that is able to respond to sudden changes in supply and demand. For example, energy demand in the UK spikes around 7am and again, between 4 and 7pm which is usually when people get up in the morning and when they return home from school or work!
Another aspect of energy security is security in the long term. With fossil fuels like oil, gas and coal, there is a limited supply and eventually these sources of energy will run out. Using renewable energy sources like wind, solar and hydro power can improve energy security in the long term.

Accessibility

Accessibility relates to a country's ability to provide universal access to reliable, affordable, and abundant energy for domestic and commercial use. The dimension captures basic access to electricity and clean cooking fuels and technologies, access to prosperity-enabling levels of energy consumption, and affordability of electricity, gas, and fuel. We need energy to live our every day lives: to heat our homes, run our cars and public transport and power the lights in buildings. It is important that the energy that we use is affordable and accessible to everyone. According to the International Energy Agency's 2020 report, solar power is the cheapest source of electricity in history, with wind power not too far behind. This is partly down to more efficient solar plants and wind turbines to allow for better energy generation. However, issues include whether they allow reliable energy provision. We can also improve energy affordability by making more energy efficient products. Gadgets that take less energy to power can help drive down energy costs by lessening demand.

Finally, please watch this 3 minute video on the energy trilemma:

https://www.youtube.com/watch?v=CI4DnLsANJM

Load libraries

```
library(quanteda)
library(quanteda.textstats)
library(quanteda.textplots)
library(stringr)
library(openxlsx)
library(ggplot2)
library(lme4)
library(MuMIn)
library(sjPlot)
library(irr)
library(DescTools)
library(lattice)
library(party)
```

Load data

Load keywords, but remove any that were suggested as a result of the analysis of the human judgements, to avoid circularity.

```
kw = read.csv("../data/LEXIS/TrilemmaKeywords.csv",stringsAsFactors = F)
kw = kw[kw$Notes != "Suggested by human judgements",]

getKeywords= function(sub){
  kx = unique(unlist(strsplit(kw[kw$Subject==sub,]$concepts,";")))
  names(kx) = kx
  return(kx)
}
accessibilityKeywords = getKeywords("Accessibility")
securityKeywords = getKeywords("Security")
sustainabilityKeywords = getKeywords("Sustainability")
```

Load human ratings of articles:

```
d4 = NULL
for(file in list.files("../data/HumanJudgements/judgements/")){
  dx = NULL
  if(grepl("xlsx",file)){
    dx = read.xlsx(paste0("../data/HumanJudgements/judgements/",file),1)
  }
  if(grepl("csv",file)){
    dx = read.csv(paste0("../data/HumanJudgements/judgements/",file),stringsAsFactors = F)
  }
  dx$participant = as.numeric(gsub("_","",substr(file,14,15)))
  d4 = rbind(d4,dx)
}
# Exclude articles rated twice
d4 = d4[!duplicated(paste(d4$participant,d4$ID)),]

#d4$text = copUK[match(d4$ID,copUK$ID),]$text
d4$totalJudgement = d4$Sustainability + d4$Security + d4$Accessibility

getProp = function(dx,measure){
```

```
  X = dx[,measure] / dx$totalJudgement
  X[dx$totalJudgement==0] = 0
  return(X)
}

d4$SustainabilityProp = getProp(d4,"Sustainability")
d4$SecurityProp = getProp(d4,"Security")
d4$AccessibilityProp = getProp(d4,"Accessibility")

d4$ID2 = paste0(d4$ID,"_",d4$participant)
d4$participant = factor(d4$participant)
```

Load reference corpus frequencies for alternative measure. The frequencies come from the SiBol Extended corpus of UK newspaper articles from the last 10 years (see Dunning, 1993; Partington, 2010), as made available on Sketch Engine (Kilgariff et al., 2014).

```
refFreqAcc = read.csv("../data/EngBroadsheetNewspaperCorpus/acc.csv",
                stringsAsFactors = F,skip=2)
refFreqSec = read.csv("../data/EngBroadsheetNewspaperCorpus/sec.csv",
                stringsAsFactors = F,skip=2)
refFreqSus = read.csv("../data/EngBroadsheetNewspaperCorpus/sus.csv",
                stringsAsFactors = F,skip=2)
```

Function to compare frequencies between two corpora, based on the G2 metric (see Rayson et al., 2004).

```
logLikelihood.G2 = function(a,b,c,d){
  # freqInCorpus1,freqInCorpus2,sizeOfCorpus1,sizeOfCorpus2
  E1 = c*(a+b) / (c+d)
  E2 = d*(a+b) / (c+d)
  G2 = 2*((a*log(a/E1)) + (b*log(b/E2)))
  G2[a==0] = NA
  return(G2)
}
```

Function to load article text and calculate frequency scores:

```
getTextFromParticipantFile = function(partNum,ids){
  fn = paste0("../data/HumanJudgements/stimuli/NewsArticles_",
              partNum,"_SENT.txt")
  tx = readLines(fn)
  tx = paste(tx,collapse="\n")
  tx = gsub("\n  [0-9][0-9]?\n","\n\n",tx)
  tx = strsplit(tx,"\n [0-9][0-9]? : COP")[[1]]
  idx = str_extract(tx,"[0-9][0-9]?_UK[0-9]+")
  idx = paste0("COP",idx)
  tx[match(ids,idx)]
}


processFile = function(d, accessibilityKW,securityKW,sustainabilityKW,
                       refFreqAcc,refFreqSec,refFreqSus){
  # Get text from file sent to participant
  d$text = ""
  for(px in unique(d$participant)){
    d[d$participant==px,]$text =
      getTextFromParticipantFile(px,d[d$participant==px,]$ID)
```

```
}
# Lower case
d$text = tolower(d$text)

# some texts need to be borrowed from other files
d[is.na(d$text),]$text = d[match(d[is.na(d$text),]$ID, d$ID),]$text

# Create corpus, tokens, freq matrix
corp = corpus(d, docid_field = "ID2",text_field = "text")
tok = tokens(corp, remove_punct = TRUE)
corpDFM = dfm(tok)
d$ArticleTotalWords = rowSums(corpDFM)

# Get frequency for one keyword
getFrequency = function(keyword){
  keyword = tolower(keyword)
  if(grepl(" ",keyword)){
    # Multi-word expression
    return(sapply(str_extract_all(d$text, keyword),length))
  }
  if(keyword %in% colnames(corpDFM)){
    return(as.vector(corpDFM[,keyword]))
  }
  return(rep(0,nrow(d)))
}

# Get score (frequency per 1000 words)
getScore = function(keywords){
  freq = sapply(keywords,getFrequency)
  return(rowSums(freq))
  #prop = 1000000 * (freq/totalWords)
  #return(prop)
}

d$CorpAccFreq = getScore(accessibilityKW)
d$CorpSecFreq = getScore(securityKW)
d$CorpSusFreq = getScore(sustainabilityKW)

d$CorpAccFreqRel = (1000000 * d$CorpAccFreq) / d$ArticleTotalWords
d$CorpSecFreqRel = (1000000 * d$CorpSecFreq) / d$ArticleTotalWords
d$CorpSusFreqRel = (1000000 * d$CorpSusFreq) / d$ArticleTotalWords

getScoreG2 = function(keywords,kFreq){
  freq = sapply(keywords,getFrequency)
  refFreq = kFreq[match(colnames(freq),kFreq$Item),]$Frequency
  refFreq[is.na(refFreq)] = 0
  G2s = sapply(1:nrow(freq),function(i){
    logLikelihood.G2(freq[i,],refFreq,
                     d$ArticleTotalWords[i],482360)
  })
  return(colMeans(G2s,na.rm = T))
}
```

```
  d$CorpSusFreqG2 = getScoreG2(sustainabilityKW,refFreqSus)

  return(d)
}
```

```
d4 = processFile(d4,
    accessibilityKeywords, securityKeywords, sustainabilityKeywords,
    refFreqAcc,refFreqSec,refFreqSus)
```

```
## Warning in readLines(fn): incomplete final line found on
## '../data/HumanJudgements/stimuli/NewsArticles_1_SENT.txt'
```

```
d4$Sustainability.scaled = d4$Sustainability/10
d4$CorpSusFreq.scaled = d4$CorpSusFreq/max(d4$CorpSusFreq)
```

Results

Agreement between human participants

Estimate inter-rater reliability using Intraclass Correlation Coefficient:

```
commonIDs = table(d4$ID)
commonIDs = names(commonIDs)[commonIDs>6]
irrx = d4[d4$ID %in% commonIDs,]
#irrx = irrx[!duplicated(paste(irrx$participant,irrx$ID)),]
irrx = irrx[order(irrx$participant,irrx$ID),]
irrx = matrix(irrx$Sustainability,
              ncol=length(unique(irrx$participant)))
icc(irrx, model = "oneway")
```

```
##  Single Score Intraclass Correlation
##
##    Model: oneway
##    Type : consistency
##
##    Subjects = 8
##      Raters = 12
##      ICC(1) = 0.515
##
##  F-Test, H0: r0 = 0 ; H1: r0 > 0
##     F(7,88) = 13.7 , p = 6.8e-12
##
##  95%-Confidence Interval for ICC Population Values:
##   0.279 < ICC < 0.826
```

The value is “fair” according to Cicchetti (1994).

Use correlation between participants as baseline.

```
parts = as.numeric(sort(unique(d4$participant)))
cors = matrix(NA,nrow=length(parts),ncol=length(parts))
rownames(cors) = parts
colnames(cors) = parts
d4 = d4[order(d4$ID),]
for(i in 1:length(parts)){
  for(j in 1:length(parts)){
    part1 = parts[i]
    part2 = parts[j]
    p1 = d4[d4$participant==part1 & d4$ID %in% commonIDs,]
    p2 = d4[d4$participant==part2 & d4$ID %in% commonIDs,]
    cors[i,j] = cor(p1$Sustainability,p2$Sustainability,
                          method = "kendall")
  }
}
diag(cors)=NA
```

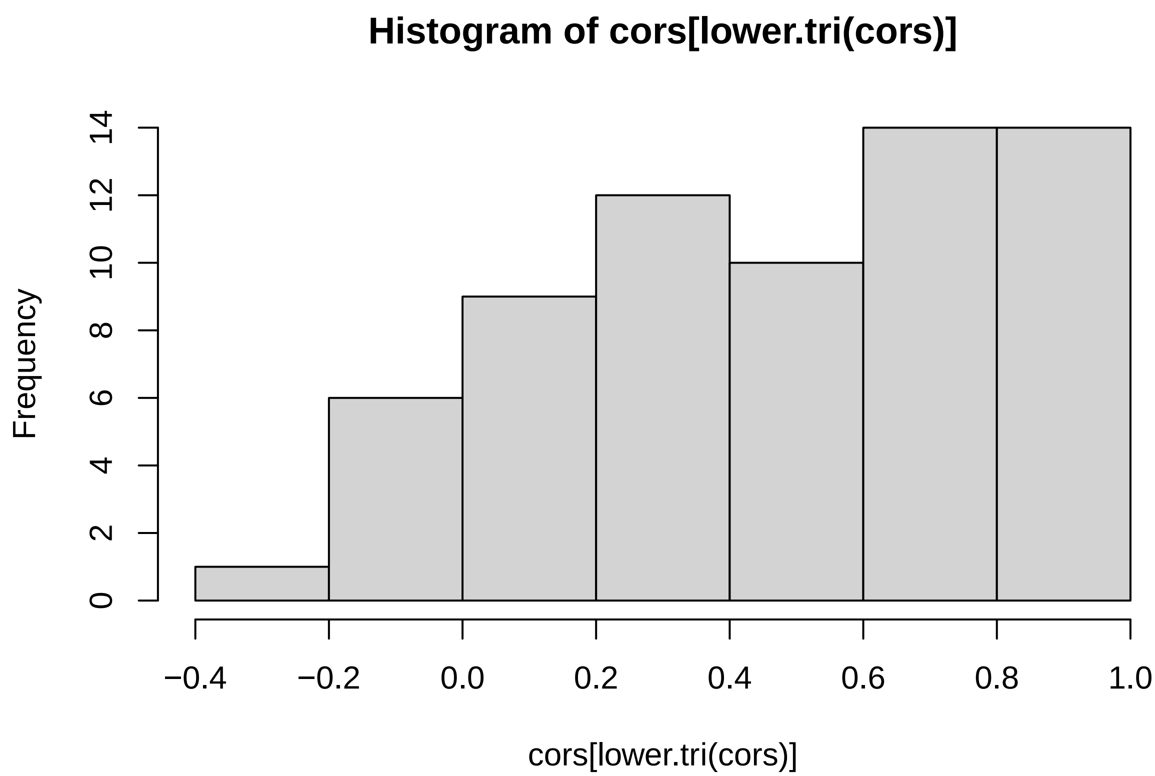
The mean correlation:

```
mean(cors[lower.tri(cors)],na.rm=T)
```

```
## [1] 0.4665243
```

There is a wide range of participant scores:

```r
hist(cors[lower.tri(cors)])
```



In particular, participant 1 seems to have very different judgements. The mean correlation increases when excluding them:

```r
meanCorBetweenHumans = mean(cors[-1,-1],na.rm=T)
meanCorBetweenHumans
```

```
## [1] 0.5508084
```

```r
cx = cors[-1,-1]
cx = cx[lower.tri(cx)]
sdCorBetweenHumans = sd(cx)
```

The ICC increases when excluding participant 1:

```r
icc(irrx[,-1], model = "oneway")
```

```
##  Single Score Intraclass Correlation
##
##    Model: oneway
##    Type : consistency
##
##    Subjects = 8
##      Raters = 11
##      ICC(1) = 0.573
##
##  F-Test, H0: r0 = 0 ; H1: r0 > 0
##     F(7,80) = 15.8 , p = 7.91e-13
```

```
##
##  95%-Confidence Interval for ICC Population Values:
##   0.331 < ICC < 0.856
```

In addition, participant 1 reported difficulties with viewing the text in the proper format. So we exclude participant 1 from the data:

```
d4 = d4[d4$participant!=1,]
d4$participant = factor(d4$participant)
```

We note that the overall average human ratings for each aspect of the trilemma reflect the broad pattern in the full data: sustainability is discussed most.

```
mean(d4$Sustainability)
```

```
## [1] 3.924658
```

```
mean(d4$Security)
```

```
## [1] 1.726027
```

```
mean(d4$Accessibility)
```

```
## [1] 1.942922
```

Agreement between human and automated measures

Raw correlation between human and automated frequency:

```
corBetweenHumansAndAuto = cor.test(d4$Sustainability, d4$CorpSusFreq,
         method = "kendall")
corBetweenHumansAndAuto
```

```
##
##  Kendall's rank correlation tau
##
## data:  d4$Sustainability and d4$CorpSusFreq
## z = 12.934, p-value < 2.2e-16
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.4529647
```

Correlation between human and relative frequency:

```
cor.test(d4$Sustainability, d4$CorpSusFreqRel,
         method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  d4$Sustainability and d4$CorpSusFreqRel
## z = 11.335, p-value < 2.2e-16
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.3866621
```

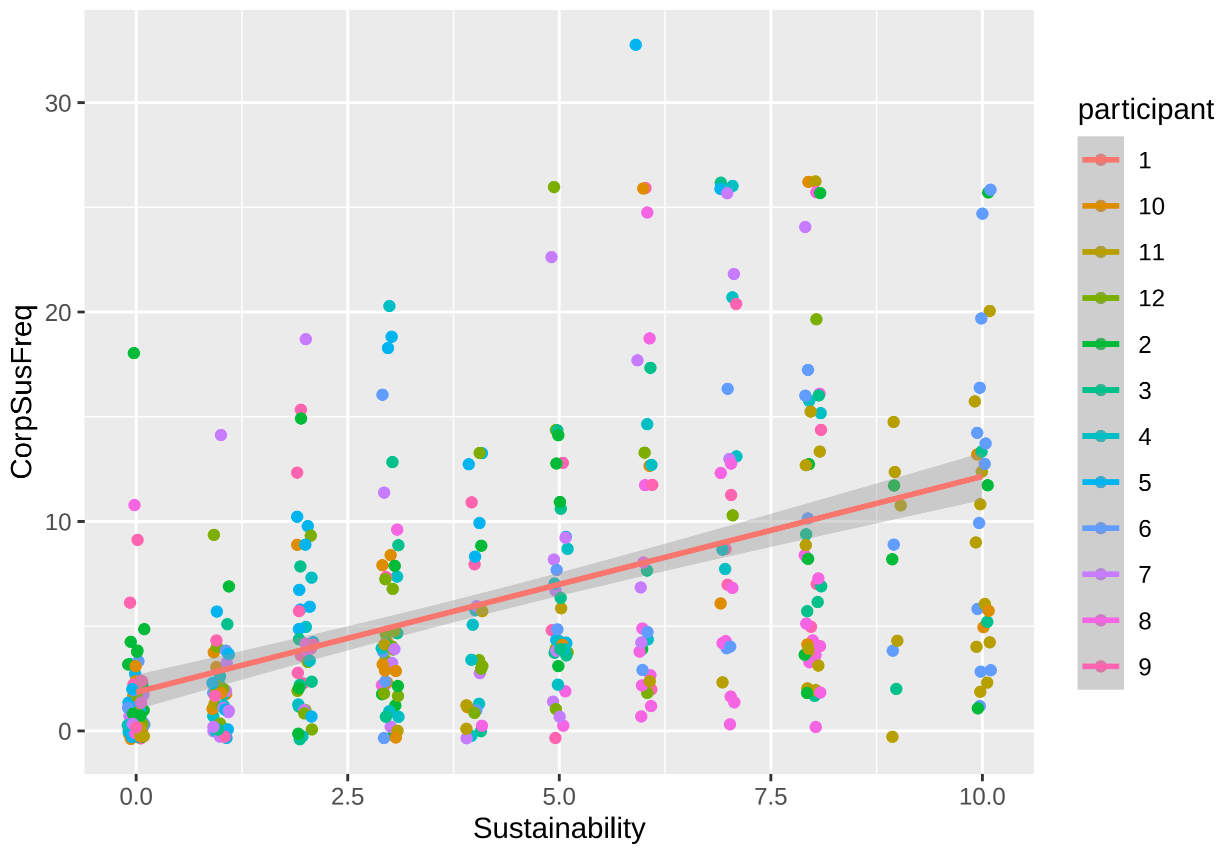
Raw correlation for the alternative measure, based on G2:

```
cor.test(d4$Sustainability, d4$CorpSusFreqG2,
         method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  d4$Sustainability and d4$CorpSusFreqG2
## z = 1.3103, p-value = 0.1901
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##        tau
## 0.04840456
```

```
ggplot(d4, aes(x=Sustainability, y= CorpSusFreq,colour=participant)) +
  geom_jitter(width=0.1) +
  geom_smooth(aes(colour="1"),method = "lm")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



Modelling the participant ratings is conceptually difficult, so we just model the frequencies using Poisson regression. We add an intercept for each participant to remove the random influence of participant baselines.

```
mx = glmer(CorpSusFreq~ Sustainability.scaled +
             (1|participant),
           family=poisson,
           data = d4)
mxSummary = summary(mx)
mxSummary
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: CorpSusFreq ~ Sustainability.scaled + (1 | participant)
##    Data: d4
##
##      AIC      BIC   logLik deviance df.resid
##   3190.8   3203.1  -1592.4   3184.8      435
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0454 -1.4613 -0.6128  0.8330 10.7880
##
## Random effects:
##  Groups      Name        Variance Std.Dev.
##  participant (Intercept) 0.08434  0.2904
## Number of obs: 438, groups:  participant, 11
```

```
##
## Fixed effects:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)            0.73618    0.09925   7.417  1.2e-13 ***
## Sustainability.scaled  2.13884    0.07523  28.430  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## Sstnblty.sc -0.426
```

The model accounts for a high proportion of the variance. The marginal effect is around 0.63, which puts an estimate for the correlation at 0.79.

```
# Correlation between observed and predicted values:
cor.test(d4$Sustainability.scaled,
         predict(mx),method="kendall")
```

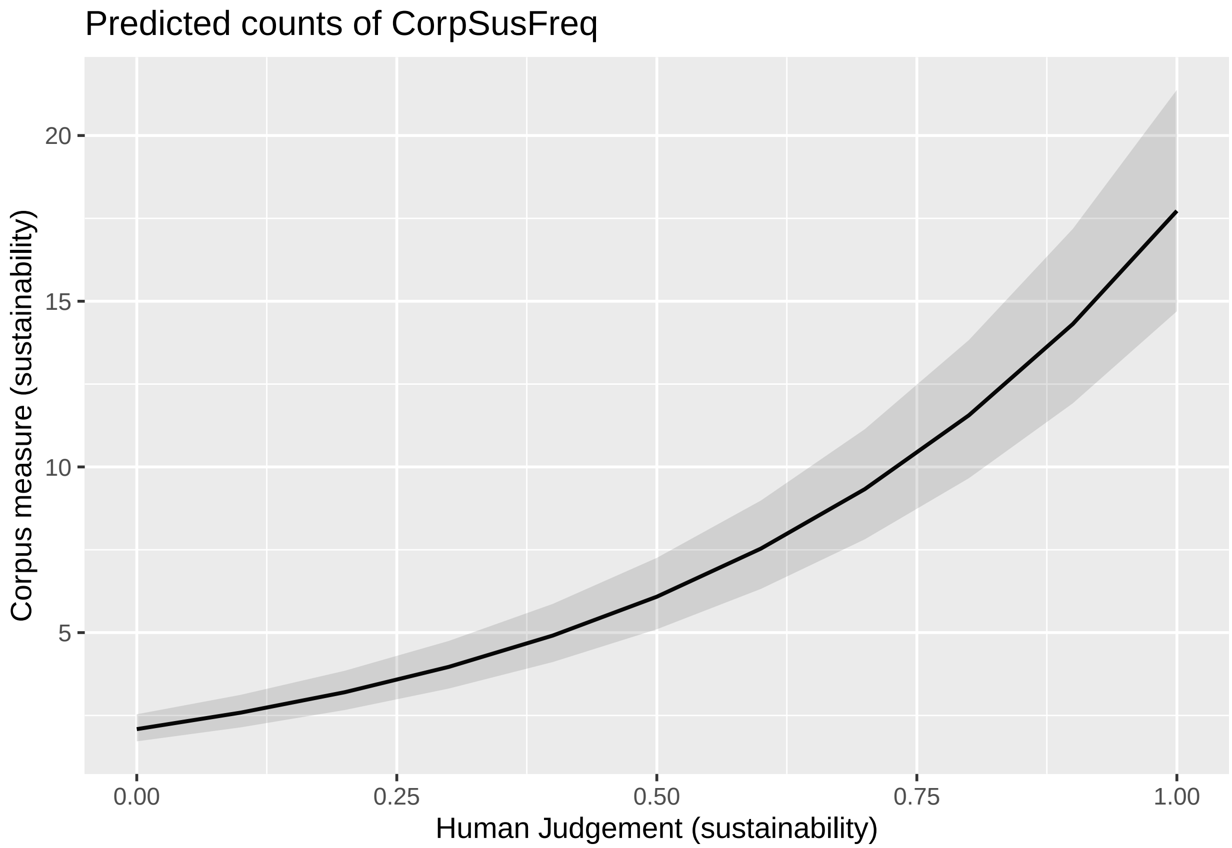
```
##
##  Kendall's rank correlation tau
##
## data:  d4$Sustainability.scaled and predict(mx)
## z = 23.184, p-value < 2.2e-16
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.7847387
```

```
# Pseudo R-squared:
mxRSq = r.squaredGLMM(mx)
mxRSq
```

```
##                 R2m       R2c
## delta     0.6508925 0.7668002
## lognormal 0.6626876 0.7806958
## trigamma  0.6375678 0.7511028
```

Clear positive relationship (it appears non-linear because of the Poisson regression):

```
plot_model(mx,"eff")[[1]] +
  xlab("Human Judgement (sustainability)") +
  ylab("Corpus measure (sustainability)")
```



Other measures:

```
cor.test(d4$Security, d4$CorpSecFreq,
         method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  d4$Security and d4$CorpSecFreq
## z = 7.1653, p-value = 7.761e-13
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.2921498
```

```
cor.test(d4$Accessibility, d4$CorpAccFreq,
         method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  d4$Accessibility and d4$CorpAccFreq
## z = 6.2559, p-value = 3.953e-10
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.2635583
```

Extra keywords

Find words that appear more frequently in texts that are accurately predicted compared with texts that are under-predicted.

```
suggestKeywords = function(d4, mx, breaks=c(-3,-2,2)){
  d4$resid = resid(mx)

  lowTexts = d4[d4$resid < breaks[1],]
  midTexts = d4[d4$resid > breaks[2] & d4$resid < breaks[3],]

  corpL = corpus(lowTexts,text_field="text")
  colloc = textstat_collocations(tokens(corpL,remove_punct = T))
  cat("Frequent Collocations in under-estimated texts:\n")
  print(head(colloc,n=20))

  kwtexts = rbind(lowTexts,midTexts)
  corp = corpus(kwtexts,text_field="text")
  toks <- tokens(corp, remove_punct = TRUE)
  dfmat <- dfm(toks)
  tstat_key <- textstat_keyness(dfmat,target = kwtexts$resid > -2)
  cat("-------------\n\n")
  cat("Keywords (frequent in well-predicted vs. under-predicted\n")
  print(head(tstat_key,n=20))
}
```

Sustainability

```
mxSus = glmer(CorpSusFreq~ Sustainability.scaled +
                (1|participant),
            family=poisson,
            data = d4)
suggestKeywords(d4,mxSus)
```

```
## Frequent Collocations in under-estimated texts:
##              collocation count count_nested length   lambda         z
## 1         climate change    66            0      2 7.975912 14.084527
## 2        paris agreement    19            0      2 8.358752 13.435205
## 3               could be    18            0      2 6.330328 13.104696
## 4                 of the    57            0      2 1.881757 11.430918
## 5                  it is    14            0      2 3.684528 11.038006
## 6                   is a    19            0      2 2.943709 10.817982
## 7           young people    10            0      2 8.261183 10.551336
## 8           under threat     6            0      2 6.951389 10.465907
## 9          world leaders     7            0      2 6.672736 10.342041
## 10          some islands     6            0      2 5.567032 10.109082
## 11            rising sea     6            0      2 7.334167  9.935094
## 12 african agriculture     5            0      2 6.659172  9.889191
## 13      commitments made     5            0      2 6.547819  9.859394
## 14                we are     8            0      2 4.134598  9.742276
## 15        united nations    14            0      2 8.791735  9.692805
## 16      development bank     6            0      2 7.782120  9.495894
## 17       pacific islands     6            0      2 6.666539  9.487467
## 18               will be     9            0      2 3.702060  9.474500
## 19               be gone     9            0      2 6.078010  9.472360
## 20     cop23 conference     6            0      2 5.680419  9.369772
## -------------
##
## Keywords (frequent in well-predicted vs. under-predicted
##       feature      chi2            p n_target n_reference
## 1       delta 11.387234 0.0007395056      180           0
## 2         had  7.904516 0.0049311534      125           0
## 3      summit  7.893970 0.0049599871      184           2
## 4   emissions  7.705038 0.0055066922      352           9
## 5    minister  7.524758 0.0060856723      119           0
## 6   president  7.309542 0.0068589416      201           3
## 7          he  7.237328 0.0071403183      343           9
## 8         his  6.292565 0.0121245447      130           1
## 9      carbon  5.896478 0.0151711934      272           7
## 10     fossil  5.816274 0.0158785479       92           0
## 11      cop21  4.912643 0.0266608094      160           3
## 12      talks  4.751832 0.0292671156      132           2
## 13        gas  4.574251 0.0324559381      129           2
## 14         us  4.529817 0.0333091878      223           6
## 15    glasgow  4.514008 0.0336183790      153           3
## 16        set  4.250655 0.0392351898       97           1
## 17    targets  4.250655 0.0392351898       97           1
## 18   expected  4.049298 0.0441897100       80           0
## 19       bonn  3.986754 0.0458593220      119           2
## 20      green  3.976989 0.0461259413      167           4
```

```r
textX = gsub("http.+ ?"," ",d4$text)
tokX = tokens(corpus(textX),remove_punct = T)
head(textstat_keyness(dfm(tokX),
        target = d4$SustainabilityProp>0.75),
     n=40)
```

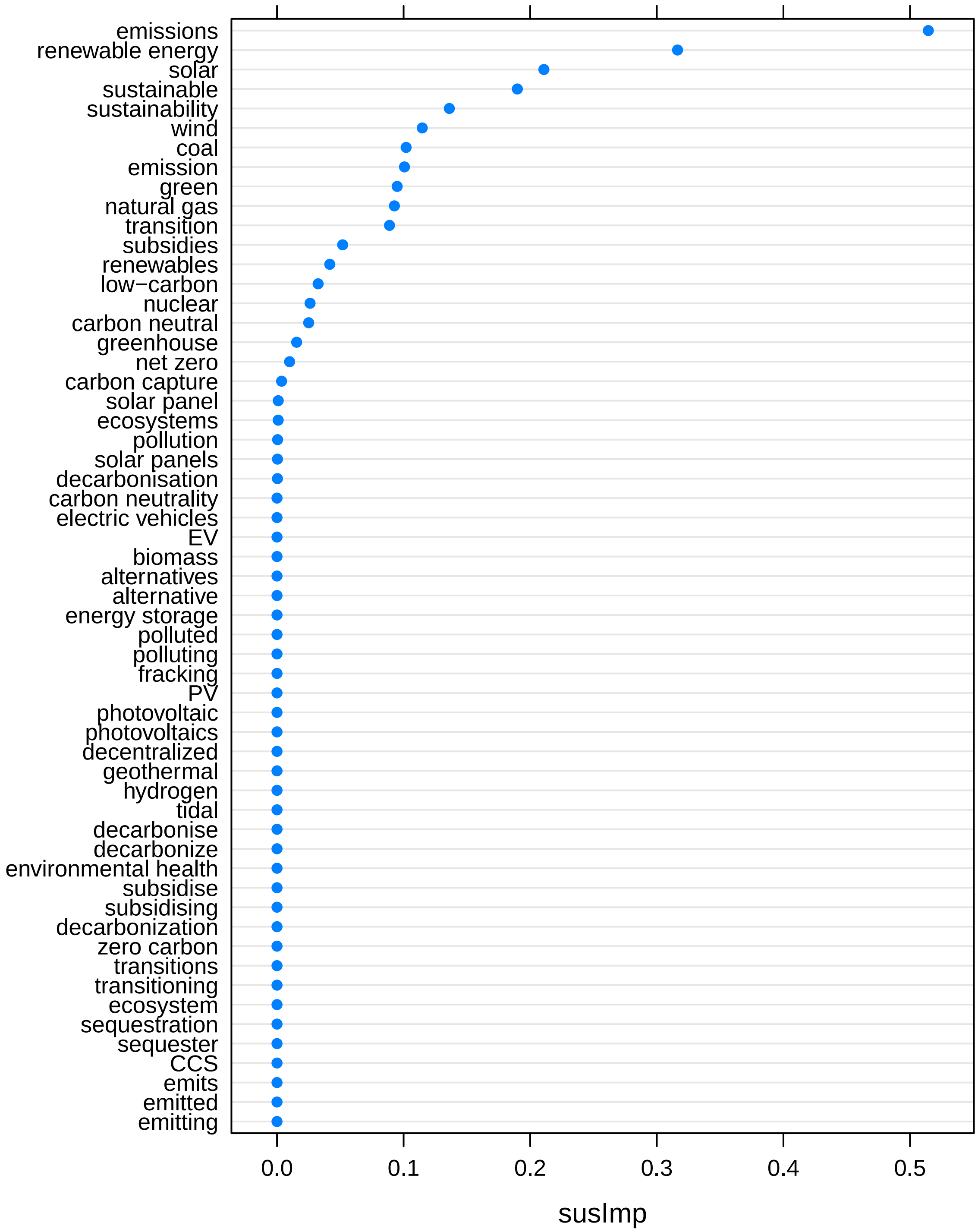
```
##          feature      chi2            p n_target n_reference
## 1          delta 145.69325 0.000000e+00      144          62
## 2          cocoa  73.72558 0.000000e+00       43           6
## 3       shipping  69.89769 1.110223e-16       35           2
## 4            imo  66.71866 3.330669e-16       30           0
## 5      buildings  49.98387 1.550204e-12       57          29
## 6    electronics  42.76906 6.159873e-11       48          24
## 7          trump  37.84806 7.647424e-10       46          25
## 8          ghana  31.63190 1.863431e-08       16           1
## 9          grand  30.43784 3.447374e-08       27          10
## 10         index  29.21544 6.476105e-08       23           7
## 11    enterprise  27.99248 1.217879e-07       14           0
## 12           ics  27.99248 1.217879e-07       14           0
## 13      organize  26.82967 2.221999e-07       24           9
## 14        palais  26.82967 2.221999e-07       24           9
## 15            us  26.54043 2.580799e-07      146         185
## 16      mondelez  25.77499 3.836269e-07       13           0
## 17     emissions  25.61034 4.177948e-07      224         322
## 18          past  25.29895 4.909764e-07       53          44
## 19 commonwealth  24.38628 7.882826e-07       14           1
## 20     reduction  24.25269 8.448886e-07       63          59
## 21         cop21  23.81060 1.062955e-06       99         114
## 22             +  22.34923 2.273047e-06       36          25
## 23       ireland  20.04635 7.558752e-06       12           1
## 24     intensity  19.95024 7.948381e-06       25          14
## 25         mills  19.13037 1.220865e-05       10           0
## 26            we  18.65835 1.563612e-05      392         663
## 27          2007  18.39194 1.798171e-05       14           4
## 28         young  18.35365 1.834670e-05       40          34
## 29     recycling  18.07969 2.118496e-05       15           5
## 30      honduras  17.88991 2.340591e-05       11           1
## 31     agreement  17.71842 2.561341e-05      232         363
## 32         paris  17.43468 2.973517e-05      282         458
## 33        france  17.07471 3.593772e-05       36          30
## 34      doconomy  16.91954 3.899812e-05        9           0
## 35    bangladesh  16.40804 5.106807e-05       21          12
## 36   secretariat  16.12988 5.914338e-05       19          10
## 37   constructed  15.97066 6.433169e-05       16           7
## 38        taipei  15.97066 6.433169e-05       16           7
## 39         wales  15.79781 7.048415e-05       49          50
## 40      aviation  15.74574 7.245099e-05       10           1
```

Test which words are contributing most to the prediction. This uses a random forest: a machine learning method of predicting a value based on the presence or absence of features.

```r
getWordImportance = function(d,measure,keywords){
  d$text = tolower(d$text)
  corp = corpus(d, docid_field = "ID2",text_field = "text")
  tok = tokens(corp, remove_punct = TRUE)
```

```r
  corpDFM = dfm(tok)
  getFrequency = function(keyword){
    keyword = tolower(keyword)
    if(grepl(" ",keyword)){
      # Multi-word expression
      return(sapply(str_extract_all(d$text, keyword),length))
    }
    if(keyword %in% colnames(corpDFM)){
      return(as.vector(corpDFM[,keyword]))
    }
    return(rep(0,nrow(d)))
  }
  freq = sapply(keywords,getFrequency)
  freq = freq[,colSums(freq)>0]
  freq = as.data.frame(freq)
  freq$SusScore = d[,measure]
  #ct = ctree(SusScore ~ ., data=freq)
  cf = cforest(SusScore ~ ., data=freq)
  imp = sort(varimp(cf))
  return(imp)
}
```

```r
susImp = getWordImportance(d4,"Sustainability",sustainabilityKeywords)
dotplot(susImp)
```


emissions
renewable energy
solar
sustainable
sustainability
wind
coal
emission
green
natural gas
transition
subsidies
renewables
low−carbon
nuclear
carbon neutral
greenhouse
net zero
carbon capture
solar panel
ecosystems
pollution
solar panels
decarbonisation
carbon neutrality
electric vehicles
EV
biomass
alternatives
alternative
energy storage
polluted
polluting
fracking
PV
photovoltaic
photovoltaics
decentralized
geothermal
hydrogen
tidal
decarbonise
decarbonize
environmental health
subsidise
subsidising
decarbonization
zero carbon
transitions
transitioning
ecosystem
sequestration
sequester
CCS
emits
emitted
emitting
0.0
0.1
0.2
0.3
0.4
0.5
susImp

Security

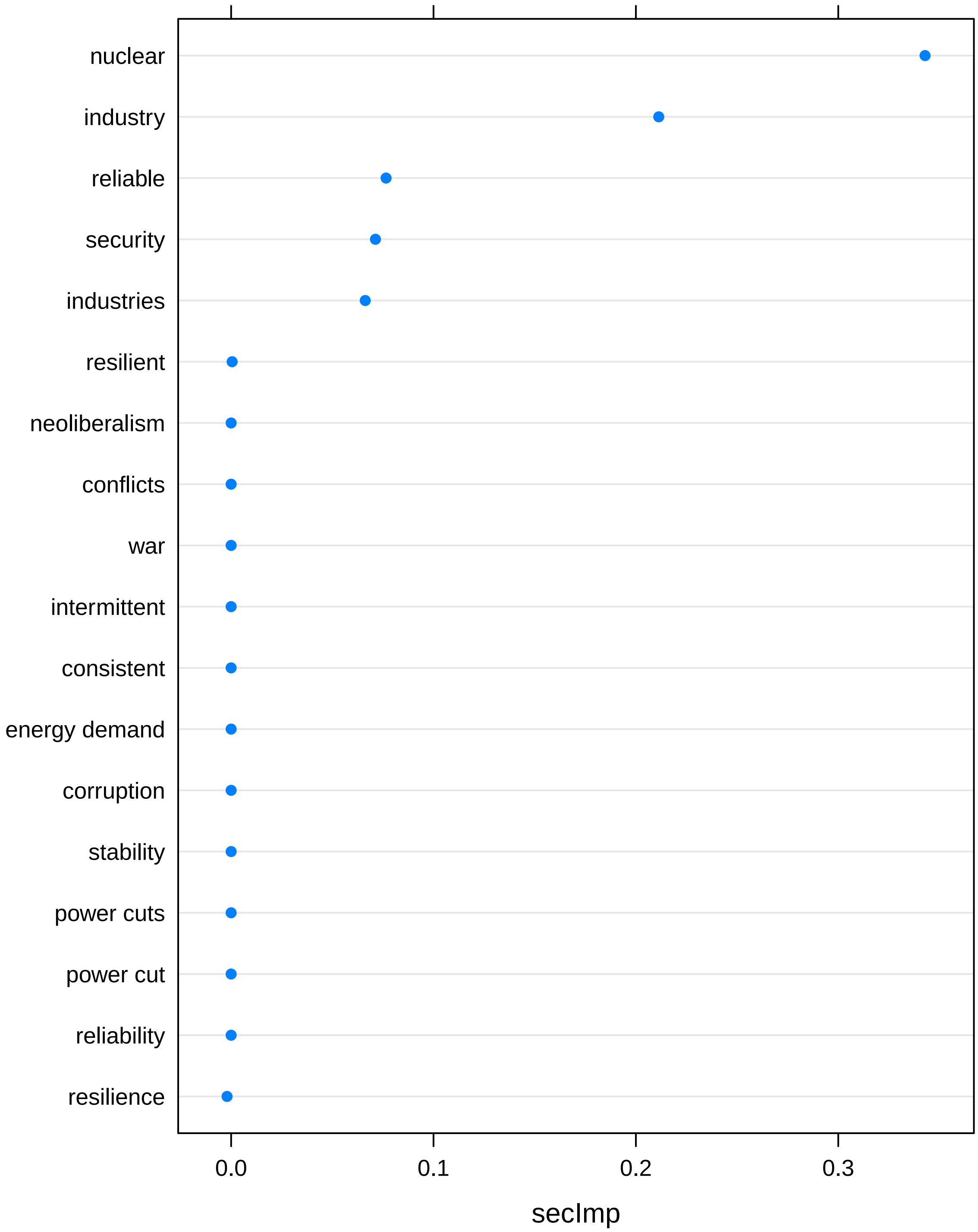
```
mxSec = glmer(CorpSecFreq~ Security +
             (1|participant),
           family=poisson,
           data = d4)
suggestKeywords(d4,mxSec)
```

```
## Frequent Collocations in under-estimated texts:
##                  collocation count count_nested length   lambda        z
## 1                  by fiscal     6            0      2 5.202981 7.829462
## 2             innovation 2050     6            0      2 6.105397 7.446849
## 3                 fiscal 2050     5            0      2 4.868873 7.326695
## 4      hitachi environmental     6            0      2 4.023047 7.241865
## 5   environmental innovation     6            0      2 5.717001 7.176986
## 6                    she had     4            0      2 4.836624 6.772715
## 7               hitachi will     5            0      2 5.400843 6.691367
## 8             climate summit     4            0      2 5.526363 6.521298
## 9               emissions at     3            0      2 5.588610 6.436343
## 10               through its     3            0      2 5.025098 6.271819
## 11  long-term environmental     4            0      2 5.200857 6.252620
## 12                  when she     4            0      2 6.138636 6.154555
## 13              thunberg had     3            0      2 4.656010 6.147780
## 14             2050 compared     4            0      2 6.037871 6.084261
## 15        resource efficient     3            0      2 7.739262 6.081055
## 16        low-carbon society     3            0      2 5.275731 5.967341
## 17            cop25 climate     3            0      2 5.184077 5.894982
## 18              resources by     3            0      2 4.309505 5.820379
## 19                  you were     2            0      2 6.043311 5.798045
## 20                 water and     5            0      2 3.449464 5.730541
## -------------
##
## Keywords (frequent in well-predicted vs. under-predicted
##          feature     chi2           p n_target n_reference
## 1            the 9.669757 0.001873261    11853          67
## 2             be 8.874407 0.002891961     1082           0
## 3      countries 5.515705 0.018846476      674           0
## 4           have 4.035226 0.044559715      718           1
## 5             to 4.019099 0.044987742     6129          36
## 6          paris 3.852225 0.049680054      695           1
## 7            are 3.550981 0.059510248     1040           3
## 8      hyperlink 3.317096 0.068562849      522           0
## 9             we 3.173368 0.074847951      987           3
## 10            he 3.050753 0.080699705      489           0
## 11            is 2.928588 0.087023575     1935           9
## 12          this 2.650841 0.103495122      912           3
## 13        change 2.540744 0.110942478     1232           5
## 14          more 2.439309 0.118328337      612           1
## 15        carbon 2.216527 0.136539896      385           0
## 16        energy 2.055875 0.151620140      560           1
## 17        global 2.048493 0.152356957      649           2
## 18         cop26 2.025559 0.154671958      361           0
## 19     agreement 2.012160 0.156043065      554           1
## 20 international 1.583462 0.208262741      305           0
```

```r
head(textstat_keyness(dfm(tokX),
        target = d4$SecurityProp>0.5),
     n=20)
```

```
##                feature      chi2            p n_target n_reference
## 1             circular 248.04498 0.000000e+00       19          10
## 2                nepal 141.55792 0.000000e+00        9           2
## 3                  adb 132.67949 0.000000e+00        7           0
## 4               status 117.21408 0.000000e+00        9           4
## 5              nuclear 111.89020 0.000000e+00       30          99
## 6             devolved 110.68688 0.000000e+00        6           0
## 7                paper  82.11393 0.000000e+00        8           6
## 8                 ldcs  72.47545 0.000000e+00        5           1
## 9          projections  72.09427 0.000000e+00        7           5
## 10                opic  66.91823 3.330669e-16        4           0
## 11               renew  52.16351 5.107026e-13        5           3
## 12            midlands  52.15289 5.134781e-13        4           1
## 13              trains  52.15289 5.134781e-13        4           1
## 14               built  50.63846 1.110445e-12       12          32
## 15              models  45.35921 1.640155e-11        6           7
## 16                auob  45.28612 1.702516e-11        3           0
## 17      climdev-africa  45.28612 1.702516e-11        3           0
## 18 continent-spanning  45.28612 1.702516e-11        3           0
## 19             engie's  45.28612 1.702516e-11        3           0
## 20            kathmandu  45.28612 1.702516e-11        3           0
```

```r
secImp = getWordImportance(d4,"Security",securityKeywords)
dotplot(secImp)
```


nuclear
industry
reliable
security
industries
resilient
neoliberalism
conflicts
war
intermittent
consistent
energy demand
corruption
stability
power cuts
power cut
reliability
resilience
0.0
0.1
0.2
0.3
secImp

Accessibility

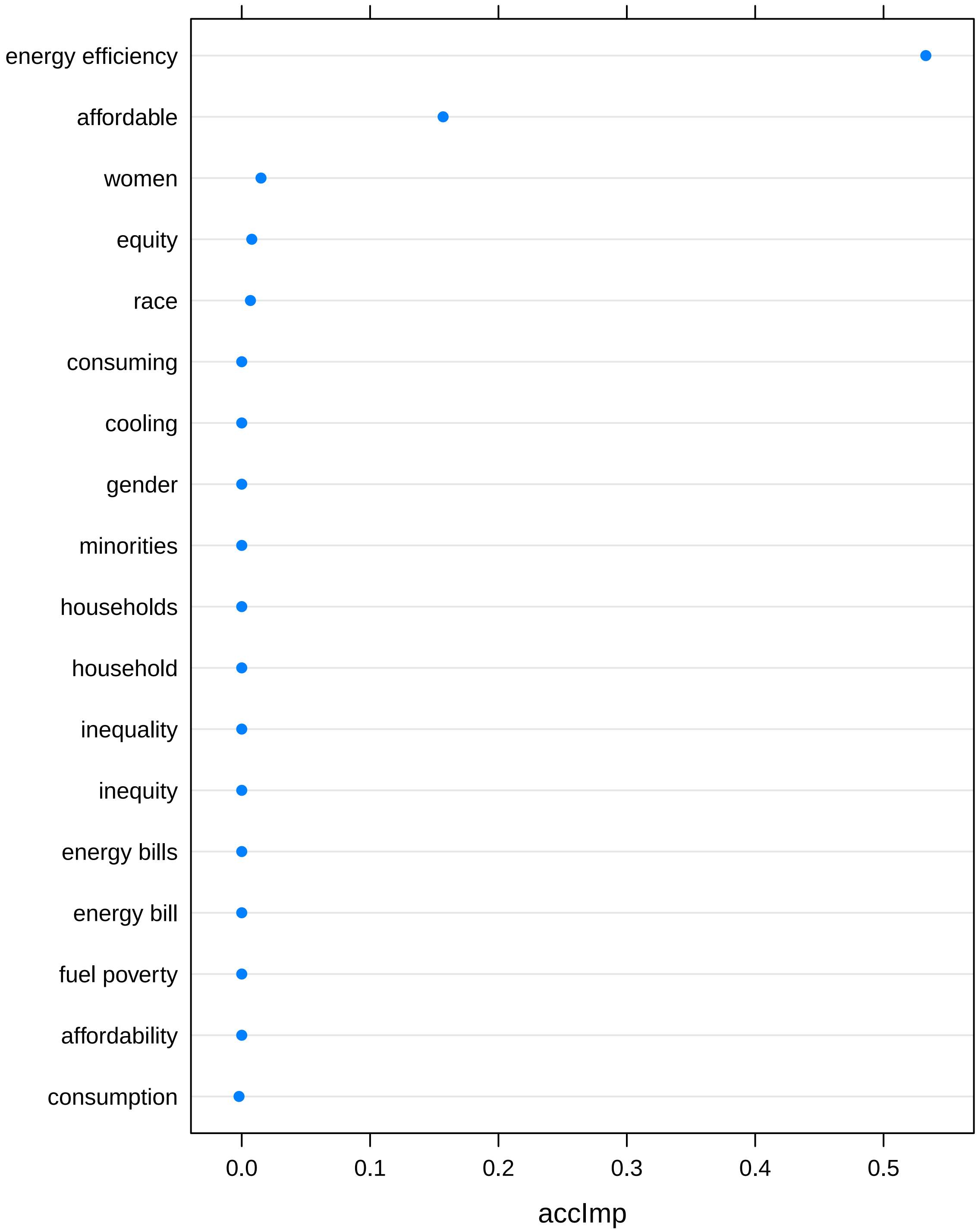
```
mxAcc = glmer(CorpAccFreq~ Accessibility +
             (1|participant),
           family=poisson,
           data = d4)
suggestKeywords(d4,mxAcc,c(-0.5, -0.5, 2))
```

```
## Frequent Collocations in under-estimated texts:
##              collocation count count_nested length   lambda        z
## 1        paris agreement   274            0      2 6.593833 59.23851
## 2         climate change   938            0      2 8.095747 53.50712
## 3                will be   250            0      2 4.371205 52.26066
## 4                 of the  1128            0      2 1.793731 48.72588
## 5         united nations   219            0      2 8.816664 44.73931
## 6              more than   109            0      2 6.506797 40.75081
## 7                 we are   144            0      2 3.880094 38.78856
## 8      change conference   138            0      2 4.034104 38.73002
## 9                  it is   170            0      2 3.562986 38.44857
## 10                at the   406            0      2 2.487489 36.66109
## 11               he said    96            0      2 4.678888 36.61609
## 12              has been    86            0      2 5.085950 36.22001
## 13             have been    84            0      2 5.100866 36.03728
## 14      carbon emissions    77            0      2 5.184023 35.76425
## 15               such as   122            0      2 5.933075 35.64732
## 16      development bank    67            0      2 6.532912 35.33985
## 17 african development    62            0      2 6.004362 34.52876
## 18            africa day    57            0      2 6.921133 34.44651
## 19                in the   725            0      2 1.501005 34.30612
## 20               we have   102            0      2 3.954082 33.82180
## -------------
##
## Keywords (frequent in well-predicted vs. under-predicted
##          feature      chi2         p n_target n_reference
## 1             we 2.5638710 0.1093306     1008           0
## 2              a 2.4120570 0.1204041     3227           6
## 3           will 1.9421674 0.1634334     1281           1
## 4             at 1.5830831 0.2083171     1163           1
## 5         energy 1.3022680 0.2537994      633           0
## 6      hyperlink 0.9833185 0.3213809      535           0
## 7      emissions 0.8683225 0.3514201      499           0
## 8            but 0.7893704 0.3742906      474           0
## 9           said 0.7414560 0.3891949      865           1
## 10           not 0.7268361 0.3939107      454           0
## 11          that 0.5930145 0.4412558     1696           4
## 12            be 0.5528068 0.4571731     1153           2
## 13        carbon 0.5434474 0.4610075      394           0
## 14            is 0.5327300 0.4654614     2001           5
## 15        reduce 0.5011790 0.4789825      145           0
## 16          very 0.5011790 0.4789825      145           0
## 17 environment 0.4977200 0.4805036      144           0
## 18             $ 0.4769666 0.4897996      138           0
## 19     companies 0.4700491 0.4929646      136           0
## 20       finance 0.4700491 0.4929646      136           0
```

```r
head(textstat_keyness(dfm(tokX),
        target = d4$AccessibilityProp>0.4),
     n=20)
```

```
##        feature     chi2            p n_target n_reference
## 1         bank 80.56251 0.000000e+00       72         124
## 2       africa 66.42959 3.330669e-16       74         147
## 3  south-south 65.09067 6.661338e-16       12           0
## 4      african 58.43350 2.098322e-14       65         129
## 5         viet 52.46493 4.379830e-13       11           1
## 6          nam 40.78493 1.699390e-10        9           1
## 7  development 39.15335 3.917914e-10       98         287
## 8  cooperation 37.43762 9.438405e-10       22          27
## 9    pneumonia 34.99601 3.303820e-09        8           1
## 10         day 30.53742 3.274893e-08       48         115
## 11        hull 29.30302 6.189919e-08        6           0
## 12     support 28.94063 7.463100e-08       60         163
## 13   indonesia 28.47772 9.478292e-08       11           7
## 14       water 28.45986 9.566155e-08       51         130
## 15 pan-african 27.95701 1.240410e-07       13          11
## 16    offshore 26.05768 3.313687e-07        8           3
## 17       ifema 25.88457 3.624555e-07       12          10
## 18   insurance 25.88457 3.624555e-07       12          10
## 19   diarrhoea 24.63703 6.920886e-07        7           2
## 20         g20 24.63703 6.920886e-07        7           2
```

```r
accImp = getWordImportance(d4,"Accessibility",accessibilityKeywords)
dotplot(accImp)
```



Summary of additional keywords

Several key terms were identified above and added to the list of key terms that go into the calculation of the final scores.

Below we re-calculate the agreement with human judgements based on the full list.

```
kwAll = read.csv("../data/LEXIS/TrilemmaKeywords.csv",stringsAsFactors = F)
getKeywords2= function(sub){
  kx = unique(unlist(strsplit(kwAll[kwAll$Subject==sub,]$concepts,";")))
  names(kx) = kx
  return(kx)
}
accessibilityKeywords2 = getKeywords2("Accessibility")
securityKeywords2 = getKeywords2("Security")
sustainabilityKeywords2 = getKeywords2("Sustainability")

d4B = processFile(d4,
    accessibilityKeywords2, securityKeywords2, sustainabilityKeywords2,
    refFreqAcc,refFreqSec,refFreqSus)
d4B$Sustainability.scaled = d4B$Sustainability/10
```

Raw correlation:

```
cor.test(d4B$Sustainability, d4B$CorpSusFreq,
         method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  d4B$Sustainability and d4B$CorpSusFreq
## z = 12.899, p-value < 2.2e-16
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.4480516
```

Model:

```
mx2 = glmer(CorpSusFreq~ Sustainability.scaled +
             (1|participant),
           family=poisson,
           data = d4B)
summary(mx2)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: CorpSusFreq ~ Sustainability.scaled + (1 | participant)
##    Data: d4B
##
##      AIC      BIC   logLik deviance df.resid
##   3150.1   3162.3  -1572.0   3144.1      435
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7889 -1.4862 -0.4909  0.9704  8.3520
##
## Random effects:
##  Groups      Name        Variance Std.Dev.
##  participant (Intercept) 0.05857  0.242
## Number of obs: 438, groups:  participant, 11
```

```
##
## Fixed effects:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)            1.08048    0.08347   12.95   <2e-16 ***
## Sustainability.scaled  1.84098    0.06704   27.46   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## Sstnblty.sc -0.434
```

```
cor.test(d4B$Sustainability.scaled,
         predict(mx2),method="kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  d4B$Sustainability.scaled and predict(mx2)
## z = 23.283, p-value < 2.2e-16
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.7881049
```

```
r.squaredGLMM(mx2)
```

```
## Warning: the null model is correct only if all variables used by the original
## model remain unchanged.
```

```
##                 R2m       R2c
## delta     0.6359932 0.7421456
## lognormal 0.6468404 0.7548032
## trigamma  0.6240068 0.7281586
```

The additions have improved the model, but only by a very small amount.

Other measures:

```
cor.test(d4B$Security, d4B$CorpSecFreq,
         method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  d4B$Security and d4B$CorpSecFreq
## z = 5.6066, p-value = 2.063e-08
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## 0.215341
```

```
cor.test(d4B$Accessibility, d4B$CorpAccFreq,
         method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  d4B$Accessibility and d4B$CorpAccFreq
```

```
## z = 9.1405, p-value < 2.2e-16
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.3610296
```

Write summary data:

```
dOut = d4[,c("ID","participant","Sustainability","Security","Accessibility",
        "CorpSusFreq","CorpSecFreq","CorpAccFreq"),]
dOut = dOut[order(dOut$participant,dOut$ID),]
write.csv(dOut,
  "../data/HumanJudgements/humanJudgementsSummary.csv",row.names=F)
```

Summary

The automated measure reliably correlates with human judgements (mixed effects model pseudo marginal R^2 = 0.651, z = 28.43, $p < 0.001$), at Kendall's tau = 0.453.

Human judgements correlated with each other at mean Kendall's tau = 0.551, sd = 0.273. Therefore, the automated judgements are within 0.36 standard deviations of human agreement.

An alternative measure of frequency based on measuring frequency in the target article over and above the relative frequency in typical news articles was also tested. However, this correlated much less well with human judgements.

References

Cicchetti, D.V., 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological assessment, 6(4), p.284.

Dunning, T.E., 1993. Accurate methods for the statistics of surprise and coincidence. Computational linguistics, 19(1), pp.61-74.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. and Suchomel, V., 2014. The Sketch Engine: ten years on. Lexicography, 1(1), pp.7-36.

Partington, A., 2010. Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS) on UK newspapers: an overview of the project. Corpora, 5(2), pp.83-108.

Rayson, P., Berridge, D. and Francis, B., 2004. Extending the Cochran rule for the comparison of word frequencies between corpora. In 7th International Conference on Statistical analysis of textual data (JADT 2004) (pp. 926-936).