well, language, human language, the bridges that help human preserve and transfer knowledge over 10000 thousand years. I always, have a high interest in this field, but never have change to dive into it.

Surprisingly, English, the most popular language in the world, with entire Western culture backup in it, there is only around 39000 common words

How rich wolfram linguistic data is

In[•]:=
```
WordList["CommonWords"]
```

Out[•]=
```
{a, aah, aardvark, aback, abacus, abaft, abalone, abandon, abandoned, abandonment, abase,
 abasement, abash, abashed, abashment, abate, abatement, abattoir, abbe, abbess, abbey,
 abbot, abbreviate, abbreviated, abbreviation, abdicate, abdication, abdomen, abdominal,
 abduct, abducting, abduction, abductor, abeam, abed, aberrant, ··· 39 104 ···, zeta, zeugma,
 zigzag, zilch, zillion, zinc, zinfandel, zing, zinger, zinnia, zip, zipper, zippy, zircon,
 zirconium, zit, zither, zloty, zodiac, zodiacal, zombie, zonal, zone, zoning, zoo, zoological,
 zoologist, zoology, zoom, zoophyte, zounds, zucchini, zwieback, zydeco, zygote, zygotic}
```

Full expression not available (original memory size: 1.8 MB)

And, you maybe find different number of words in difference sources of dictionaries, but I believe around 150000 words of English we know.

In[•]:=
```
WordData[]
```

Out[•]=
```
{0, 1, 10, 100, 1000, 10000, 100000, 1000000, 1000000000, 1000000000000, 1000th, 100th, 101,
 101st, 105, 105th, 10-membered, 10th, 11, 110, 110th, 115, 115th, 11 November, 11-plus,
 11th, 12, 120, 120th, 125, 125th, 12th, ··· 149 127 ···, Zygocactus truncatus, zygodactyl,
 zygodactyl foot, zygoma, zygomatic, zygomatic arch, zygomatic bone, zygomatic process,
 zygomorphic, zygomorphous, Zygomycetes, Zygomycota, Zygomycotina, Zygophyllaceae,
 Zygophyllum, Zygophyllum fabago, Zygoptera, zygospore, zygote, zygotene, zygotic, Zyloprim,
 zymase, zymogen, zymoid, zymology, zymolysis, zymolytic, zymosis, zymotic, zymurgy, Zyrian}
```

Full expression not available (original memory size: 7.2 MB)

You can check more at “https://en.wikipedia.org/wiki/List_of_dictionaries_by_number_of_-words”. Another high credential dictionary is Oxford Advanced Learner ‘s Dict, have 145000 words. We can expect that Wolfram team spent time curated their English word data. You can check their sources here https://reference.wolfram.com/language/note/WordDataSourceInformation.html

In[•]:=
```
WordData["love", "Properties"] // Multicolumn[#, 4] &
```

Out[•]=

AmericanSpelling	ConsequencesTerms	MorphologicalDerivatives	SubsetTerms
Antonyms	Definitions	MorphologicalSource	SupersetTerms
BaseAdjective	DerivedAdjectives	NarrowerTerms	Synonyms
BaseForm	DerivedAdverbs	PartsOfSpeech	UsageField
BaseNoun	EntailedTerms	PartTerms	UsageType
BritishSpelling	Examples	PhoneticForm	WholeTerms
BroaderTerms	GeographicDomain	PorterStem	WordNetID
CausesTerms	Hyphenation	SentenceFrames	
CompositeTerms	InflectedForms	SimilarAdjectives	
ConceptWeight	MaterialTerms	SimilarVerbs	

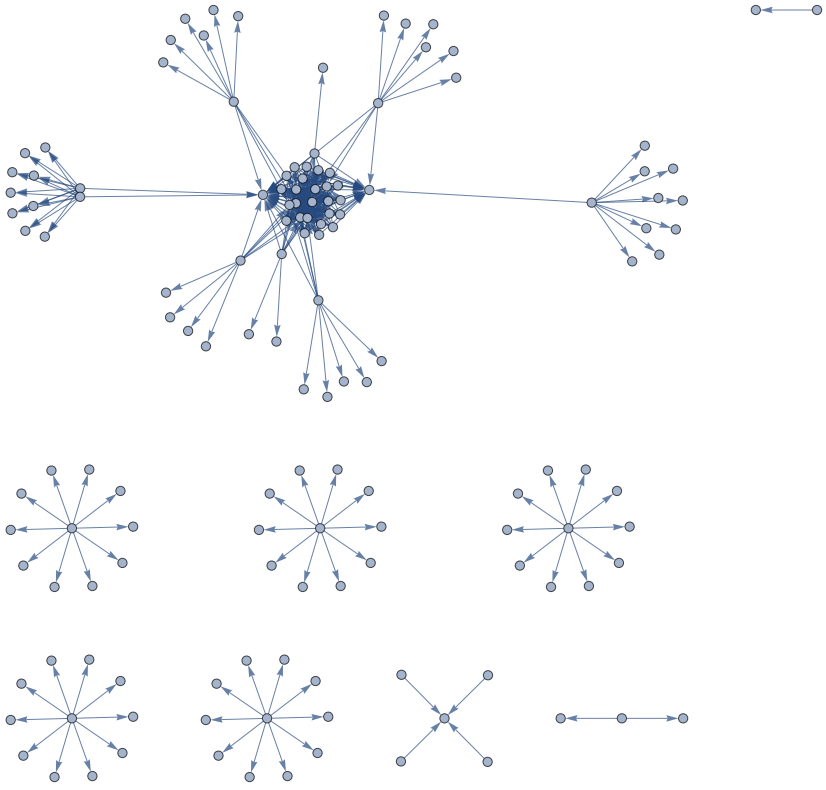
Let see, we have 37 properties, let's check how rich words data from Wolfram

```
generateLabels[graph_] := Module[
  {keys = Keys@graph, values = Values@graph},
  labels = keys ⋃ values;
  # → Placed[#, Tooltip] & /@ labels


 ]
```

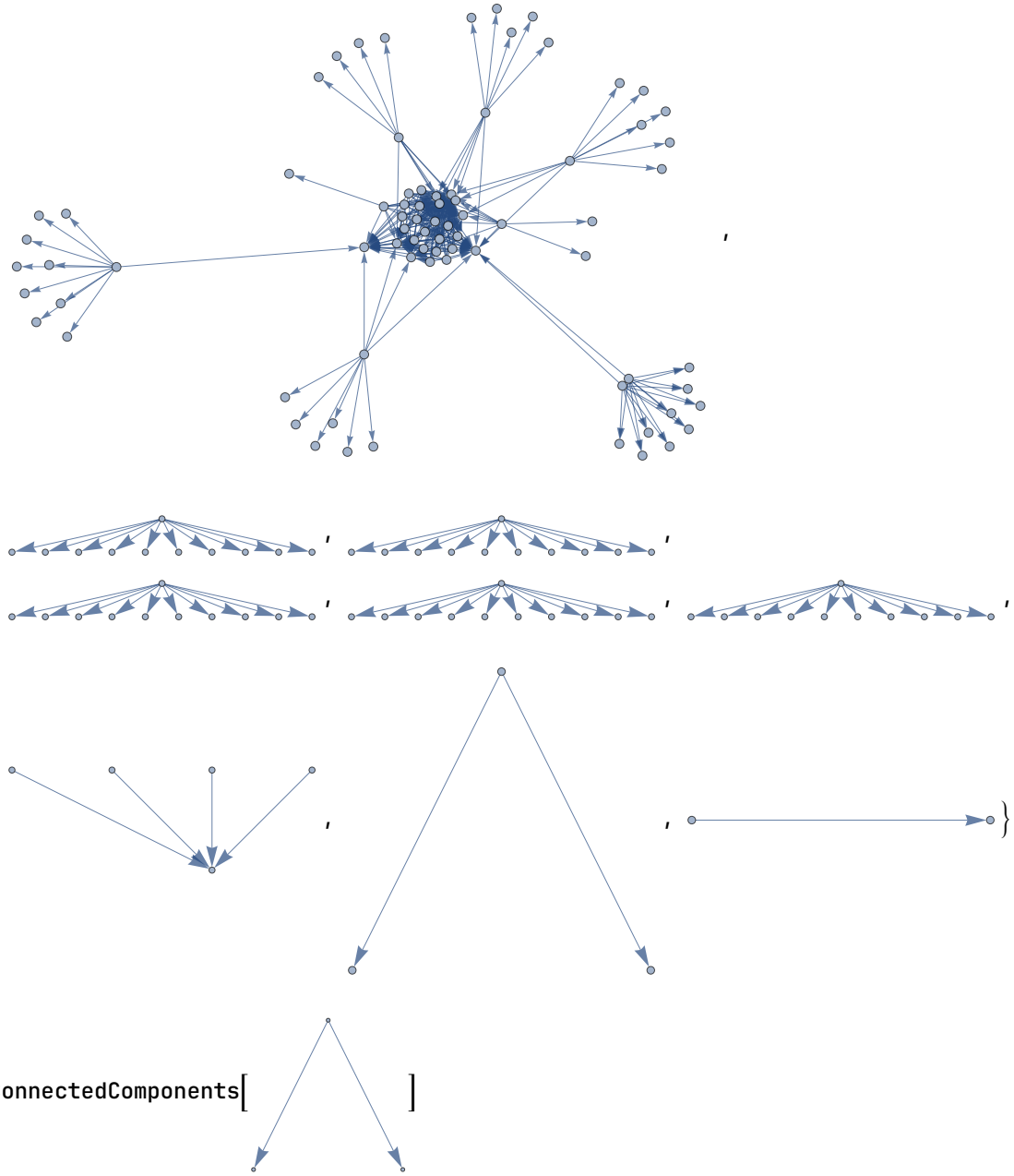
In[•]:=
```
graphLove =
 Thread /@ (# → WordData["love", #] & /@ WordData["love", "Properties"]) // Flatten //
  Graph[#, VertexLabels → generateLabels@#] &
```

Out[•]=



In[•]:=
```
graphLove // WeaklyConnectedGraphComponents
```

Out[•]=



In[•]:=
```
ConnectedComponents[ ]
```


Out[•]=
```
{{Verb}, {Noun}, {PartsOfSpeech}}
```

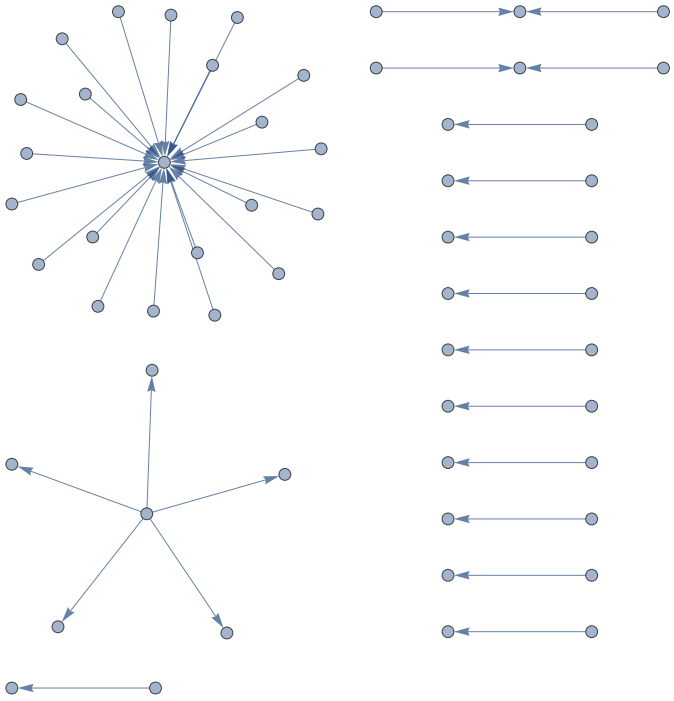
In[•]:=
```
{VertexCount@graphLove, EdgeCount@graphLove}
```

Out[•]=
```
{143, 317}
```

Very rich and careful curated, a simple “love” word return a graph with 317 edges and 143 vertexes. Linguistic data from Wolfram is more than a rich dictionary

```
In[•]:= randomWord = RandomWord[1][[1]]
        Thread /@ (# → WordData[randomWord, #] & /@ WordData[randomWord, "Properties"]) //
          Flatten // Graph[#, VertexLabels → generateLabels@#] &
Out[•]= exploratory
```

Out[•]=



Scratchpad

```
In[•]:= ({"telecommuting", "Noun"} → {}) // Last
Out[•]= {}

In[•]:= TypeOf[123]
Out[•]= TypeSpecifier[Integer64]

In[•]:= WordData["telecommuting", "BaseForm"]
Out[•]= {$Failed}

In[•]:= WordData["telecommuting", "PorterStem"]
Out[•]= telecommut

In[•]:= WordData["telecommuting", "PhoneticForm"]
Out[•]= teluhkuhmy'ooting

In[•]:= WordData["telecommuting", "InflectedForms"]
Out[•]= {{telecommuting, Noun} → {telecommutings}}
```

```
In[•]:= EntityPrefetch["Word"]

In[•]:= WordData["telecommuting", "BaseForm"]
Out[•]= {$Failed}

In[•]:= WordData[All, "Preload"]
Out[•]= True

In[•]:= PacletDataRebuild[]

In[•]:= Entity["Word", #] & /@ RandomWord[2] // AbsoluteTiming
Out[•]= {0.017011, {bowdlerize, escapement}}

In[•]:= #["BaseForm"] & /@ {bowdlerize WORD, escapement WORD} // AbsoluteTiming
Out[•]= {4.14233, {{{bowdlerize, Verb}}, {{escapement, Noun}}}}

In[•]:= Information[escapement WORD]
```

Out[•]=

Entity
escapement WORD
Canonical Name escapement
Full Dataset...

```
Entity["Word", "crosshatch"][]
Out[•]= crosshatch

In[•]:= crosshatch WORD["BaseForm"]
Out[•]= {{crosshatch, Noun}, {crosshatch, Verb}}

In[•]:= findAvailablePropertiesOfWord @"telecommuting"
```

First: Nonatomic expression expected at position 1 in First[$Failed].

```
Out[•]= {BroaderTerms → {employment, work}}

In[•]:= WordData["credential", "Synonyms"]
Out[•]= {{credential, Noun} → {certificate, certification, credentials}}

In[157]:=
SetDirectory["~/nhannht-projects/nature"];

In[•]:= NotebookSave[EvaluationNotebook[], FileNameJoin[{Directory[], "humanlanguage.nb"}]]
```

In[158]:=
```
VerminExportKeepSyntaxHighLight[]
```

In[159]:=
```
Export[FileNameJoin[{Directory[], "humanlanguage.pdf"}], EvaluationNotebook[]]
```