

## Human Rights

- Right to life or Right to belief (not a right that all humans have. Societies exist where certain beliefs are criminalized)
- Not a philosophical debate to be debated. Collectively defined by the UN after the war. So, uniform across those countries. Now, whether they enforce those rights and to what extent depends on the country. But the logical artifact is uniform throughout.
- Human rights are actionable. Because countries have some or the other legal instruments for that. Countries might choose to ignore the laws but they do exist.
- Problematic concept:
  - o How they are defined has been contested – what right does/doesn't appear on the list must have been heavily debated. Eg – (joke eg) why not a human right to Pizza but a right to life? There must be a logic. Who defined that distinction? (real eg) Right to free news – info should flow all over the world; sounded very correct (even to the US and the USSR both) but there was one country that pointed out that there was something wrong – India: Most news agencies were American. Concretely – *American* news should be available to all and that might be biased. Doordarshan will probably not be as important as CNN.
  - o **Positive and negative rights** – Negative rights are rights where you are not given anything extra but where an existing condition is protected - right to free speech/life. Positive – Right to education/food; any right that needs welfare. Human rights were biased towards negative rights and away from positive ones. *“Positive rights require welfare, so they are not rights”* – biased towards rich people.
  - o For free speech, is the free speech of a rich person owning the news channels equal to that of an average person?
  - o These worldviews were representative of US and USSR but it's more fundamental.
  - o Do you think Polio vaccines should be a human right? In a world with only negative human rights, 100s of millions of Indians would have suffered from Polio. It should be a human right. But if that is true, the human rights concept is flawed because polio vaccines are a positive right.
  - o ^Country wide debate when the rights were being defined. USSR – R&D should be distributed. US – Private property of the innovators, cannot be shared. Ended with knowledge being shared to 3<sup>rd</sup> world countries and so India was also able to make vaccines.
  - o If your criteria is bare minimum and inclusive of only negative rights, you have ignored the reality of money. Negative rights are meaningless to the poor. Right to speech meaningless if you can't speak (if the whole media is against you). **“Structural violence”** – when you are facing the system. In a system where men are treated better than women, you will face violence if you are born as a woman. No one is going out of their way to is but just because of your existence. This was not taken into account into the funda of human rights.
  - o So they replaced human rights with basic law – a list of **basic legal guarantees that society gives you** rather than a **list of rights that nature gives you**. Some countries call their constitution the basic law. Many countries have the basic law instead of human rights.

- One is to place rights on god/nature, other is on constitutional morality and third way beyond this is on structural violence. Because constitutional morality is individual based and is blind to structure. So, in India for example, poor people have extra rights which go beyond the individual level of positive/negative rights.
- Rights don't need someone else's recognition. Basic law does. Seems like a downside because at the end of the day, the State only will provide you with anything more than the negative rights.

### Short history

- People started questioning religion during the period of Enlightenment. "Give us proof". People started to look at cause and effect and questioning reality.
- Magna Carta – Great Charter. Before this, laws existed but were about duty. "You do this or we'll kill you". Magna Carta had laws that applied to the government too. Charles the first was declared the traitor and was killed by the Parliament.
- English Bill of rights – later became constitution.
- French declaration – human rights (after revolution, 1789) and since then we all have been plagiarizing.
- UN Charter (1945)
- UDHR (1948)
- However, are these rights being preserved? We are seeing 2 wars where these are being violated.
- Right against refoulement – all humans have the right to a country. Massively violated. Rohingya migrants in India, Migrants in America, etc.

Right now, what AI does externally can be governed best through/by human rights. Easy to measure – what AI does should not violate human rights. But shortcomings of human rights (not including positive rights and recognizing structural violence) apply here as well. Limitations of human rights to govern AI – don't look at socio-economic factors and structural violence.

### Word embeddings

- NLP is about language and words. ML uses maths and ML techniques use stats.
- How do we go from words to math?
  - Cat – 001 | Bat – 010 | Hat – 100 (vectors without interconnection)
  - What if they are related to each other? Eg cat's and dog's locations are closer to each other than, let's say cat's and hat's. Can we use vectors related to each other? Physical location of points can be close (lesser Euclidian distance) for words with semantic similarities.
- Word embeddings work on semantic similarity.
- How to recommend top 10 sad movies? Go to IMDB and steal top 1000 movies. Every movie description will have keywords. Make an average vector for each movie by combining the vectors for all words in that movie's description. For every movie find distance from the word "sad". The set of the 10 movies with the least distance is the answer!
- CBOW (continuous bag of words): 2 layered neural network. Take all unique words. 1 random vector for every word so you have a (random) numerical description of each word. Now we are only dealing with numbers. Now for every single word in Wikipedia – 5 ahead, 5 behind – what

will be the word in the middle? For every word, you'll have a unique mid-layer (hidden layer?) – that will be the word embedding you're trying to find.

- Skip-gram – only difference is instead of trying to predict the middle given the neighbors, it's the reverse.
- Sources on slides for where to get the word embeddings from.
  - o Word2vec (2013) – just google word2vec github for the clones.  
Application: distance.exe – In command prompt, also need to give binary name of the file (GoogleNews\_vectors....)
  - o Second link - where you can download the word vectors from.
  - o Third and fourth – Stanford's word vectors; slightly better. Trained on diff data sets.  
High number of dimensions too.
- Can convert words of the English language to numbers/vectors through which you can find similar words.
- Twig : Branch :: ? : Hand  
Answer: Fingers.  
Paris – France + India = Delhi

-----