

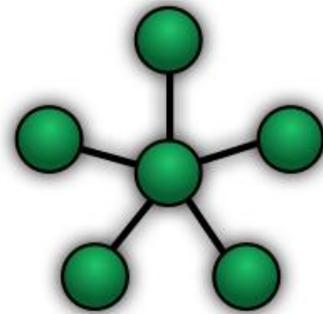
# Social Media Governance

K.S. Park

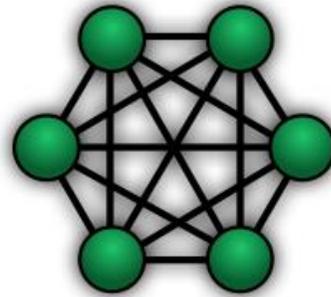
[kyungsinpark@korea.ac.kr](mailto:kyungsinpark@korea.ac.kr)

Open Net Korea/

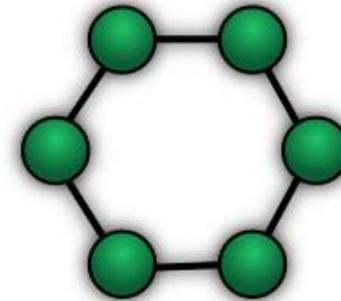
# Invention called "Internet"



Star



Fully Connected



Ring

Crowdsourcing delivery of all messages

→ All nodes talking to all other nodes without gatekeepers

→ Power of mass communication made available to all including powerless individuals

# Types of Intermediaries

- Internet service providers (ISP)?
- Webpage hosts → PLATFORMS → SOCIAL MEDIA
- Email service providers
- Search engines

# Unique Features of Social Media

- Social as in “socialization”
- Followers, subscription, retweets, “friends”
- Not expression-based but relationship-based: automated timeline curating → similar people sharing content among themselves → echo chamber and filter bubble
- Intensification of speech – “going viral”
- Calls for more governance for hate speech, fake news, and other dangerous speech

# Questions of governance

- What are the contents to be regulated by state?
- How should the contents be regulated ty state? How does it co-relate with Intermediary Liability Safe Harbor?
- Should social media be regulated like broadcasting? If not, should it still be regulated like public utilities?
- What are the contents to be regulated by social media?

# Canonical View of Freedom of speech and "False News" Regulation 1

- Speech is interactive, i.e., causes effect only through processing of the receivers → The harms caused cannot be attributed solely to the speaker or to the content because they are mediated.
- → Only speech likely to cause specific harms can be regulated (US "clear and present danger" test; EU "necessary for a democratic society" test)

# Canonical View of Freedom of speech and "False News" Regulation 2

- → Permissible speech regulations:
  - defamation (3P shunning the subject)
  - fraud(taking away of the listener's property)
  - copyright (depriving the author's market),
  - bomb hoax law(public facilities) (causing panic, "verbal act")
  - perjury (misleading on-going judicial fact-finding)
  - forgery (abusing people's trust in documents to harm them)
  - child pornography (harm in production caused by demand),
  - OBSCENITY: A BIG EXCEPTION!
  - hate speech regulation (but only majority's hate speech on minority, not the other way around b/c only the former is likely to cause the intended result)
- How about "false news" regulation? - NO specific harm anticipated (different from defamation which has a specific victim) → ergo, not consistent with human rights/constitution → also, history of being abused by authoritarian Gs for suppressing truthful dissidence

# International human rights standards clearly established

- R v. Zundel (Canada, 1992): false news regulation unconstitutional
- Chavanduka & Choto (Zimbabwe, 2000): President Mugabe's persecution of 2 reporters on "spreading falsity to cause alarm"
- Minerva case (Korea, 2010): A blogger charged with spreading false information about SKG's exchange rate policies
- Then, how to fight false news? → (1) media consumer literacy (e.g., distinguishing facts from opinion); (2) enhancing quality of professional media to "crowd out" truth; (3) government opening up more data which can be used to arrest spread of false news; the latter 2 facilitating **citizen fact-checking**
- **Marcelo Mendoza Study on Twitter on 2010 Chile earthquake: proven self-corrective capacity**

# Value of Inaccuracy for Democratic Society

*(Zundel)*

- "Should an activist be prevented from saying "the rainforest of British Columbia is being destroyed" because she fears criminal prosecution for spreading "false news" in the event that scientists conclude and a jury accepts that the statement is false and that it is likely to cause mischief to the British Columbia forest industry?"
- "Should a concerned citizen fear prosecution for stating in the course of political debate that a nuclear power plant in her neighbourhood "is destroying the health of the children living nearby" for fear that scientific studies will later show that the injury was minimal?"
- "Should a medical professional be precluded from describing an outbreak of meningitis as an epidemic for fear that a government or private organization will conclude and a jury accept that his statement is a deliberate assertion of a false fact?"
- "Should a member of an ethnic minority whose brethren are being persecuted abroad be prevented from stating that the government has systematically ignored his compatriots' plight?"

# Value of intentional lies (*Zundel*)

- “Exaggeration -- even clear falsification -- may arguably serve useful social purposes linked to the values underlying freedom of expression.
- “A person fighting cruelty against animals may knowingly cite false statistics in pursuit of his or her beliefs and with the purpose of communicating a more fundamental message, e.g., 'cruelty to animals is increasing and must be stopped'.
- “A doctor, in order to persuade people to be inoculated against a burgeoning epidemic, may exaggerate the number or geographical location of persons potentially infected with the virus.
- “An artist, for artistic purposes, may make a statement that a particular society considers both an assertion of fact and a manifestly deliberate lie; consider the case of Salman Rushdie's *Satanic Verses*, viewed by many Muslim societies as perpetrating deliberate lies against the Prophet.
- “All of this expression arguably has intrinsic value in fostering political participation and individual self-fulfilment.”

# Revival of fake news

- BuzzFeed 1: : [fake news](#) gone viral (e.g., Pope Endorses Trump) beyond real news.
- BuzzFeed 2: 40% of Trump voters believing in Democrats' child sex slave ring." 36% believing in Kenyan birth of Obama
- ➔ **"Fake news is affecting world history!!!"**
- "Fake News = NOT false news but news from FAKE SITES. Fake news is the offsprings of digitalization.
- "'METANESS' about fake news. The gravamen of fake news is not that the news is fake but the source is fake (i.e. decorated like legit news site). The harm does not come from the fact that people believe the story but they believe (incorrectly) that it was carried by reputable media.
- "➔ once believed to have been picked up by reputable media, goes viral again not because people believe the story but people believe the fact of coverage by reputable media. ➔ that alone does the magic e.g., casting a cloud of doubt on Hilary's candidacy

# My answer: Are fake news really a problem?

1. Do we know whether the stories were believed by people who shared them on Facebook? Maybe, fake news were shared just for fun not because the substance were believed. Look at [Fake sites](#) like WorldPoliticus.com, ABCNews.com.co. NOT distinguishable from supermarket tabloids ([The National Enquirer](#), [Star](#), [The Globe](#), [National Examiner](#)) or "red tops" – (e.g. [Alien Endorses Trump](#)). Will we regulate tabloids as well?

[2. Harmful controversies believed by people \(i.e. Obama's Kenyan birth\) ARE NOT FAKE NEWS](#) shared through social media but false information shared by POLITICIANS.

# Regulatory risk is greater the other side of equation: How much are we promoting truth?

- Truth defamation laws suppress truth, make it difficult to fight false news. Truth being a qualified defense is not enough. Requirement of public interest has chilling effect.
- Out of 30 or so “major” countries surveyed, Norway, Netherland, Denmark, India, Brazil, Swiss, Israel, Japan, Korea, Canada, Finland, Hungary, Italy (criminal) have truth defamation laws.
- criminal defamation chilling free media
- RTBF, another way of suppressing truth

# Way forward: BUILD AND EXPAND RESERVOIR OF TRUTH

- Abolishing criminal defamation – UK, Kenya, Zimbabwe, South Africa on a roll (includes abolition of truth defamation)
- UN Human Rights Committee in 2015 November: “Truth should be an absolute defense. Public Interest should not be a requirement” on Korea
- Moderating RTBF with privacy-based themes – “journalistic activities” exception not enough, **also unreasonably discriminating citizen journalism** notwithstanding GDPR’s broad definition (i.e., sharing info to the public)
- Also, good ole’ responses such as media literacy, more open gov data, high quality professional journalism

# Hate speech

- Hate speech is not hateful speech
- Hate speech first defined in 1948 Universal Declaration of Human Rights: speech inciting violence and discrimination against race, religion, nationality – relationship to Nazi atrocities in WWII
- Issues
  - Hate speech against majority?
  - Mental distress part of physical effect (suppressing free speech)?

# Defining hate speech

- “dirty words” vs. “speech showing hatred against others” vs. “speech inciting hatred against others”
- Is hatred always bad? E.g. “I hate racists f\_ \_ kers.”
- What hate speech is bad? → speech with a clear and present danger of causing external harm → speech against minority in a majority-dominated society where tendencies to discriminate can be easily triggered with mere words
- Also, origin of hate speech – Universal Declaration of Human Rights (UDHR) a response to racial discrimination and genocide → “no more genocide, no more genocide-triggering words” → hate speech regulation as anti-discrimination law

# Controversies about hate speech regulation 1

- Verbal discrimination enough? Or words inciting physical, financial, and institutional discrimination?
- Hatred-inciting vs. discrimination-inciting, violence-inciting
- Insulting/deprecating minorities is hate speech? (“Homosexuals are unethical and unnatural”) Or should it be more? (“Deport all homosexuals”).

# Controversies 2

- How about speech against “majority”?

e.g. “I hate angry white males oppressing us. They should burn in hell.”

# Controversies 3

- Can 'mental damage' qualify as the harm justifying hate speech regulation?
- Can 'silencing the weak' qualify as the harm justifying hate speech regulation?

# Controversies 4

- Criminal punishment
- Should hate speech be criminally punished? Or civil damages enough?
- If civil damages enough, how about just **TAKING DOWN**?

# Regulating Fakenews and Hate Speech, Lesson 1

- Internet is NOT A STADIUM. → “Putting something on Internet does not mean “publicly disclosed” → Don’t worry. People must choose from billions and billions of material. **Internet is not one huge OPEN SPACE but an aggregate of billions of rooms.** → When right-wing pastor in North Carolina burns Quoran, it caused a lot of violence in Middle East. → ho is there to know that? - **Building a Wall of Non-Recognition?**

# Lesson 2.

- Power of Internet – extremely distributed communication network → everyone's ability to speak to everyone else simultaneously without anyone's approval cf. newspaper, TV → embraces everyone's desire → Number of page views, number of clicks, why important? In microeconomics, utility achieved
- **Lesson:** Should be ready to live with *at least some* unlawful contents temporarily.

**Lesson 3:** Internet is catch-all communication that allows people to let steam off when mainstream media and politicians fail them. Internet replaces the instantaneous scribbles on bathroom stools and concrete walls. “Cold Medium v. Hot Medium” (Marshall McLuhan, *Understanding Media*, 1964)

# Various ways of regulating social media

- platform neutrality (i.e., broadcasting) – ex ante regulation
- Platform-specific laws imposing intermediary liability (Germany 2017, Australia 2018)
- general torts liability
- Mandatory identity verification

# Intermediary Liability

- When should intermediaries be held liable for “aiding and abetting” online illegalities?
- Only for known illegal contents, why? → if not, GENERAL MONITORING or Prior Censorship → Solution: safe harbor and other liability exemptions → notice-to-notice vs. noticed-and-takedown
- EU e-Commerce Directive, US DMCA/Section 230, Japan
- Manila Principles on Intermediary Liability
- How about for known contents but illegality is not known? – mandatory takedown vs. special liability vs. general tort liability, e.g., German Social Network Act 2017, Australia’s new Law 2018  
→ Should be careful about platform’s tendencies to err on the side of taking down as opposed to keeping it up

# Self-governance

- Facebook's Oversight Board v. Germany's Co-regulation model
- Oversight Board:
  - Standard: international human rights? Or freedom to innovate and compete with different standards
  - Picked by FB? – self-perpetuating?
  - Paid for by FB? – scalability
  - Regional specialization vs. moral relativism