Green Machine Translation – Sustainability & Accessibility

- Professor Sharon O’Brien and Professor Pat Brereton
- School of Applied Language and Intercultural Studies / School of Communications
- Dublin City University

GREEN SCENT
SMART CITIZEN EDUCATION FOR A GREEN FUTURE

TRANS MEDIA CATALONIA

GDA
Green Digital Accessibility

Context

- Increasing technologisation is leading to increased energy consumption
- Raises questions of climate impact and climate justice
- Friction between accessibility and sustainability
- Virtue signalling as first response
- What can the Language Technology Sector learn from the Eco-Cinema Movement?

GREEN SCENT
SMART CITIZEN EDUCATION FOR A GREEN FUTURE

TRANS MEDIA CATALONIA

GDA
Green Digital Accessibility

The ‘Material’ Turn in Eco-Cinema

- Screen industries have embraced ‘sustainability’ for three decades
- But, is it tokenistic, without industry-level changes? (Vaughan 2022)
- Increasing awareness of:
  - The material fact that film and other media are made from natural resources, energies, people
  - Need for collaboration with local stakeholders, environmental organisations and media companies to reduce environmental impact

- Vaughan, Hunter (2022) “Policy Approaches to Green Film Practices: Local Solutions for a Planetary Problem” in Kääpä & Vaughan (Eds)

The Deep Learning Turn

- Significant increase in:
  - Number of MT engines being produced
  - Number of languages being served
  - Number of users

Benefits of the DL Turn in MT

- Faster access to translated information
- Higher productivity through post-editing
- Access to ‘gisting’ feature
- Increased emphasis on digital linguistic resources

Ongoing Challenges with MT

- Fluency blindness
- Higher expectations on productivity
- Lower prices
- Sense of commodification of work
- Lack of MT literacy
- Reproduction of bias
- **Environmental Impact**

Move from CPUs to GPUs

- GPUs more energy demanding the CPUs (Shterionov & Vanmassenhove 2022)

MT R&D Demands Energy

- For the NLP community, training, intervention (tuning) and retraining cycles are the norm
- Competitions – measured in very small increments of BLEU scores – are also the norm
  - "exponentially larger models are being trained for often small gains in performance"

    (Shterionov and Vanmassenhove 2022)
- To date: little discussion about the environmental impact of such competitions

Example ‘Costs’

- DeepL: set up ‘supercomputer’ in Iceland for MT due to its ‘naturally cool climate and abundant energy’ (Slator 2021)
- Training the BERT language model is equivalent to a trans-American flight in terms of carbon emissions (Strubell et al. 2019)
- Training language models for some language pairs could be less ‘costly’ than others (Yusuf et al. 2021)

Signs of Greater Awareness (1)

- Promotion of ‘Green AI’ over ‘Red AI’ (Schwartz et al. 2020)
- First SustaiNLP workshop held in 2020
- Provision of tools to calculate environmental costs of NLP systems (Shterionov & Vanmassenhove 2022)
  - Machine Learning Emissions Calculator

Signs of Greater Awareness (2)

- Increased experimentation to calculate costs, e.g. from Shterionov & Vanmassenhove 2022
  - One experiment -> 2,500 kg of CO2
  - GPU located in Ireland had lower carbon impact than on in the Netherlands

Awareness in Translation Studies

- Growing criticism
- Sustainability largely considered from two perspectives:
  - (1) sustainability of professional practices in an increasingly technologized environment and
  - (2) the moral relationship between human beings and non-human entities. (Kenny et al. 2020)
- The Eco-Translation Turn (Cronin 2013, 2017)
  - Modern translation technology practices are deeply unsustainable
  - Should we consider a ‘Slow Language’ movement; a move from high-tech to low-tech? (Cronin 2017)

GREEN SCENT
SMART CITIZEN EDUCATION FOR A GREEN FUTURE

TRANS MEDIA CATALONIA

GDA
Green Digital Accessibility

The Dilemma: Accessibility vs Sustainability

- Improved MT means better access to information in many languages
- Especially important in situations where people are vulnerable
- Language access: portrayed as a human right (Greenwood et al 2017)
- Increased access to educational resources
- Increased focus on digitalisation of ‘smaller languages’
- Thus: to ‘slow it down’ – how will this impact negatively on these positive aspects?

Lessons from Eco-Cinema?

- Push for ‘Eco-media literacy’ -> ‘Eco-langtech literacy’? (Balance between sustainability/accessibility)
- Virtue signalling is not very helpful -> Look for solutions, don’t just highlight problems
- ScreenGreening – LangTech Greening
- Carbon Impact Statements in papers, yes, but avoid tokenism and ‘greenwashing’
- Focus on green growth and circular economy -> Smarter re-use of language data, sharing of GPUs, collaborative initiatives for low-tech MT applications for low-resource languages
- The power of storytelling to get audiences on board -> Storytelling to get developers and users on board?

GREEN SCENT
SMART CITIZEN EDUCATION FOR A GREEN FUTURE

TRANS MEDIA CATALONIA

GDA
Green Digital Accessibility