

the purpose. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of NIST, or the U.S. Government.

REFERENCES

- [1] Paola Cascante-Bonilla, Kalpathy Sitaraman, Mengjia Luo, and Vicente Ordonez. 2019. Moviescope: Large-scale Analysis of Movies using Multiple Modalities. *arXiv preprint arXiv:1908.03180* (2019).
- [2] Keith Curtis and George Awad. 2020 (accessed August 26, 2020). *DVU Challenge*. <https://drive.google.com/drive/folders/1q1Ca0aFjrF9tB8hsw-mr19d4tzy5wlpZ>
- [3] Keith Curtis, George Awad, Shahzad Rajput, and Ian Soboroff. 2020. HLVU: A New Challenge to Test Deep Understanding of Movies the Way Humans do. In *Proceedings of the 2020 International Conference on Multimedia Retrieval*. 355–361.
- [4] Jeremy Debattista, Fahim A Salim, Fasih Haider, Clare Conran, Owen Conlan, Keith Curtis, Wang Wei, Ademar Crotti Junior, and Declan O’Sullivan. 2018. Expressing Multimedia Content Using Semantics—A Vision. In *2018 IEEE 12th International Conference on Semantic Computing (ICSC)*. IEEE, 302–303.
- [5] Anna Rohrbach and Jae Sung Park. 2019. Large Scale Movie Description Challenge (LSMDC) 2019. <https://sites.google.com/site/describingmovies/lsmdc-2019>, Last accessed on 2019-11-06.
- [6] Makarand Tapaswi, Yukun Zhu, Rainer Stiefelbogen, Antonio Torralba, Raquel Urtasun, and Sanja Fidler. 2016. Movieqa: Understanding stories in movies through question-answering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4631–4640.