

- Reed, S., Zolna, K., Parisotto, E., Colmenarejo, S. G., Novikov, A., Barth-marion, G., Giménez, M., Sulsky, Y., Kay, J., Springenberg, J. T., et al. A generalist agent. *TMLR*.
- Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. High-resolution image synthesis with latent diffusion models. In *CVPR*, 2022.
- Schmidhuber, J. On learning to think: Algorithmic information theory for novel combinations of reinforcement learning controllers and recurrent neural world models. *arXiv preprint arXiv:1511.09249*, 2015.
- Srinivasan, K., Raman, K., Chen, J., Bendersky, M., and Najork, M. Wit: Wikipedia-based image text dataset for multimodal multilingual machine learning. In *SIGIR*, 2021.
- Srivastava, R. K., Greff, K., and Schmidhuber, J. Highway networks. *arXiv preprint arXiv:1505.00387*, 2015.
- Tahmasebzadeh, G., Kacupaj, E., Müller-Budack, E., Hakimov, S., Lehmann, J., and Ewerth, R. Geowine: Geolocation based wiki, image, news and event retrieval. In *SIGIR*, 2021.
- Wang, D., Wen, Z., Tong, H., Lin, C.-Y., Song, C., and Barabási, A.-L. Information spreading in context. In *WWW*, 2011.
- Wang, Z., Shan, X., Zhang, X., and Yang, J. N24news: A new dataset for multimodal news classification. In *LREC*, 2022.
- Yang, Z., Lin, Z., Kang, P., Lv, J., Li, Q., and Liu, W. Learning shared semantic space with correlation alignment for cross-modal event retrieval. *ACM TOMM*, 16(1):1–22, 2020.
- Zhang, S., Liang, Y., Gong, M., Jiang, D., and Duan, N. Multi-view document representation learning for open-domain dense retrieval. In *ACL*, 2022.