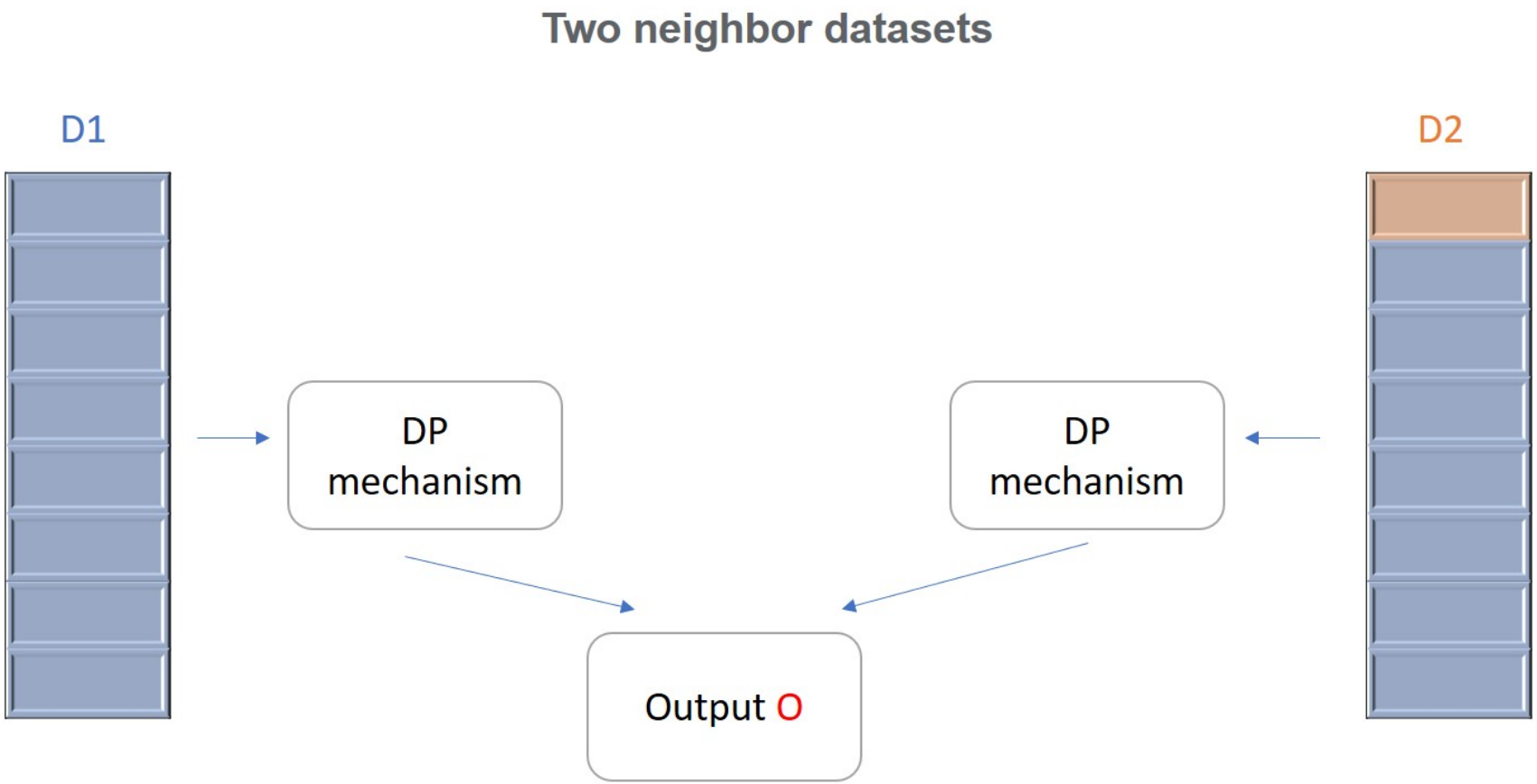
Transfer Learning In Differential Privacy's Hybrid-Model

Refael Kohen*, Or Sheffet*

*Faculty of Engineering - Bar-Ilan University

International Conference on Machine Learning (ICML) - 2022

Differential Privacy (Dwark et al., 2006)



For any set of outputs:

$$\Pr(A(D1) \in S) \leq e^{\epsilon} \Pr(A(D2) \in S) + \delta$$

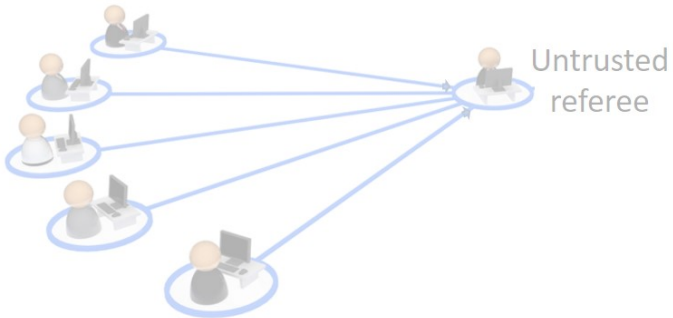
Curator and Local Models of DP

Curator model of DP:

Full access by the curator

Local model of DP:

The messages sent by the local-agents preserve DP



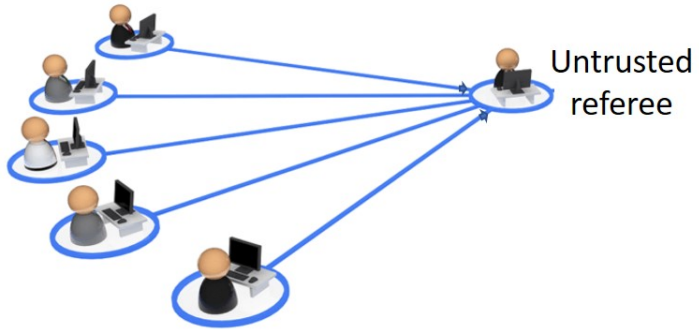
Curator and Local Models of DP

Curator model of DP:

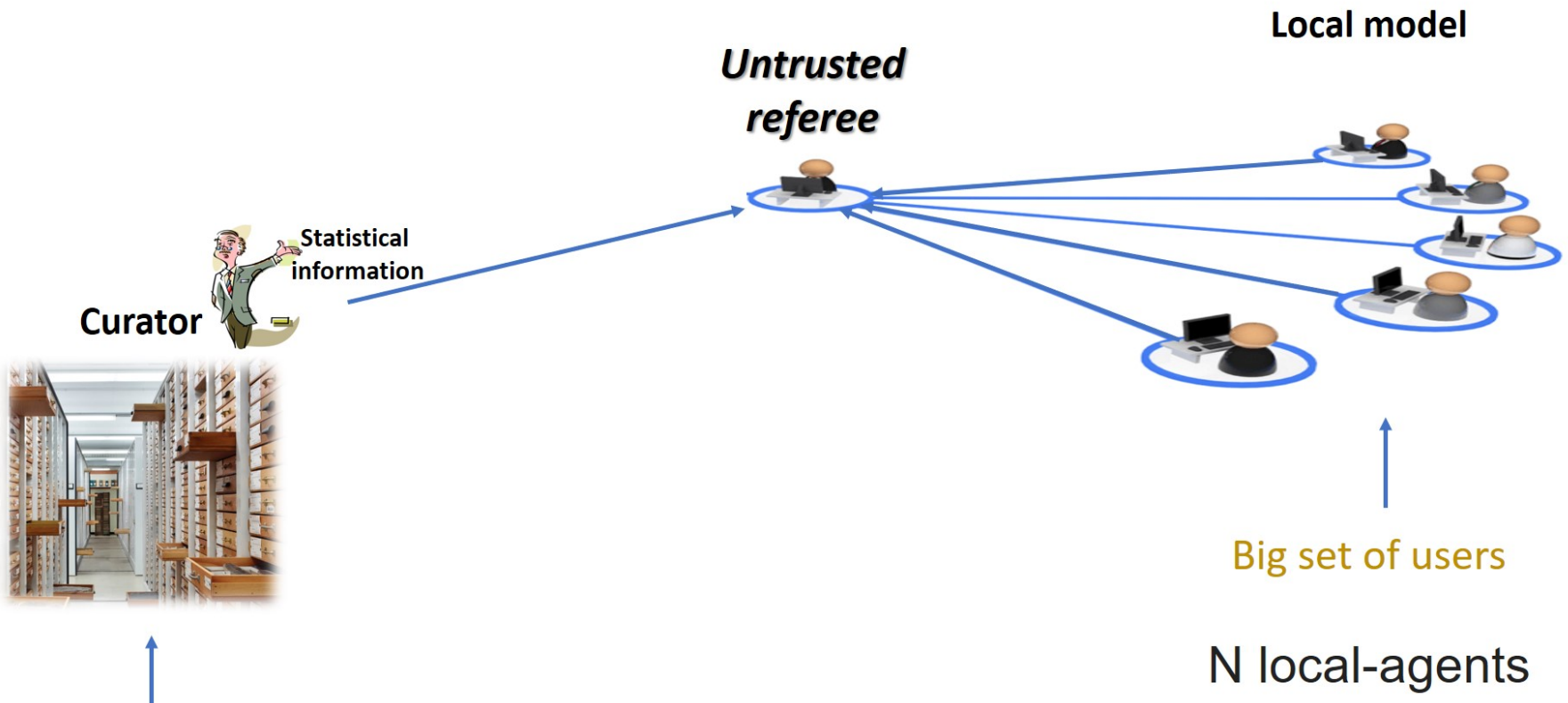
Full access by the curator

Local model of DP:

The messages sent by the local-agents preserve DP



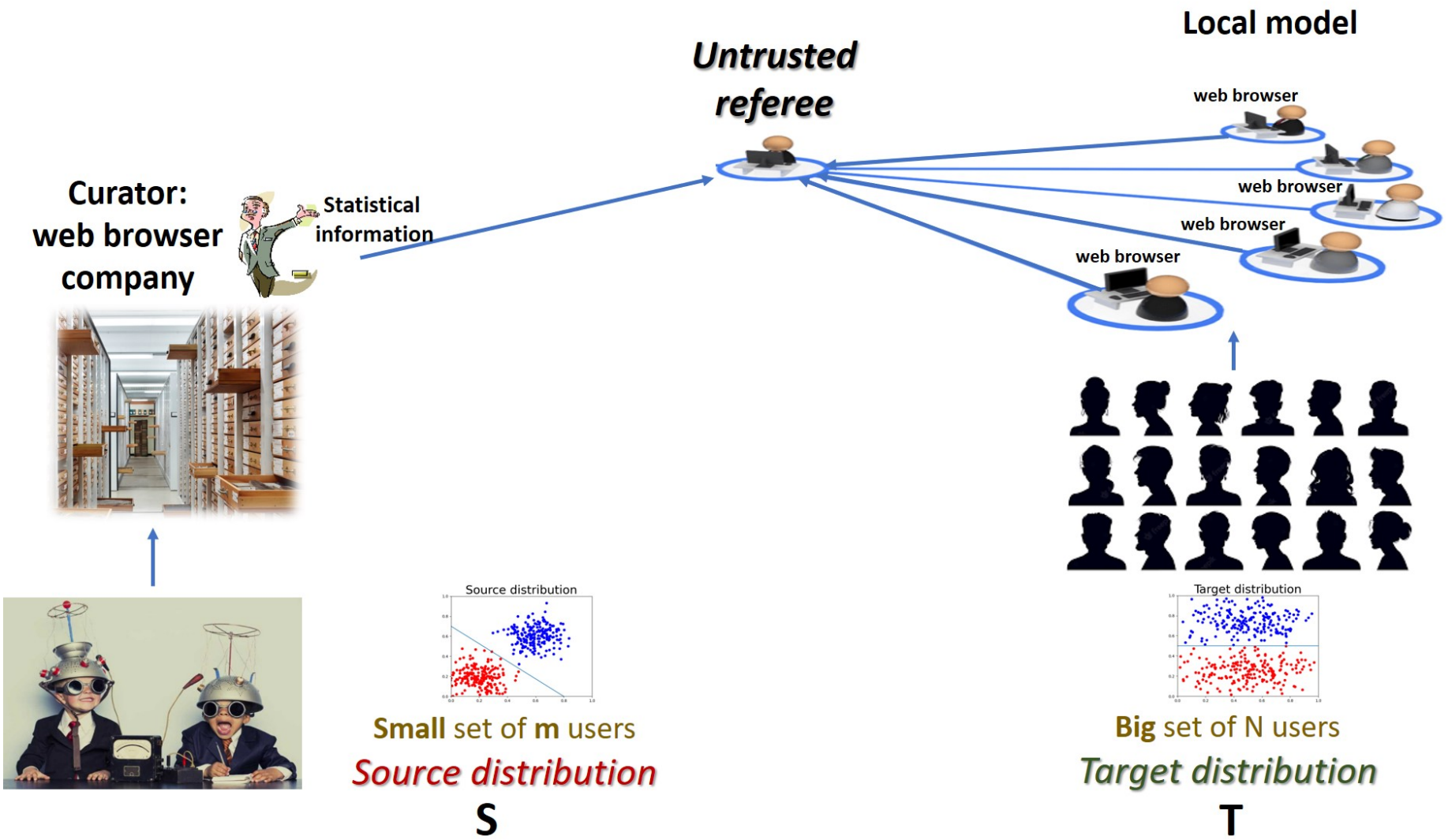
The Hybrid Model of DP (Beimel et al., 2019)



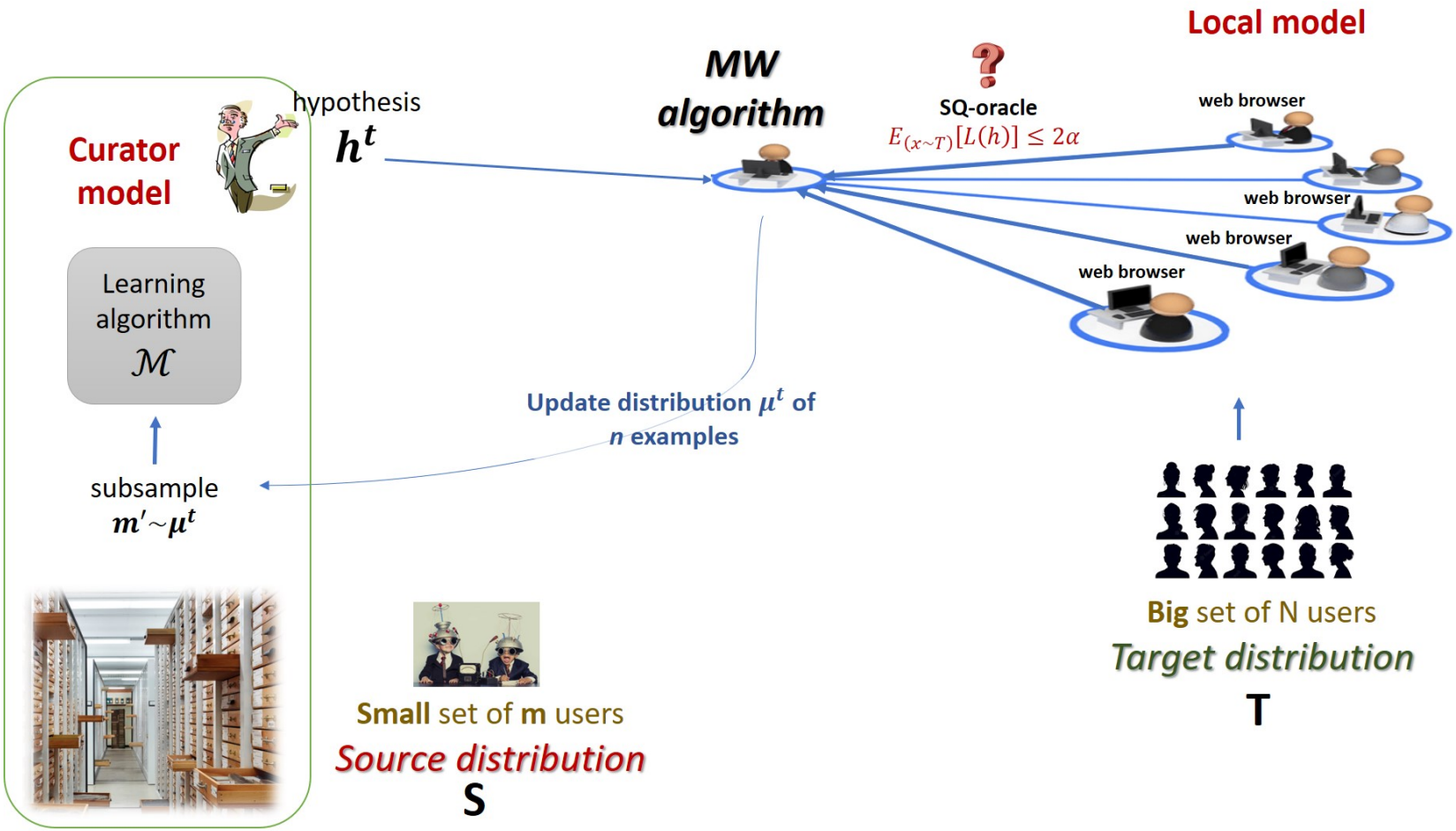
Avent et al., BLENDER: Enabling local search with a hybrid differential privacy model.

Beimel, A., Korolova, A., Nissim, K., Sheffet, O., and Stemmer, U. The power of synergy in differential privacy: Combining a small curator with local randomizers.

Hybrid Model - Transfer Learning

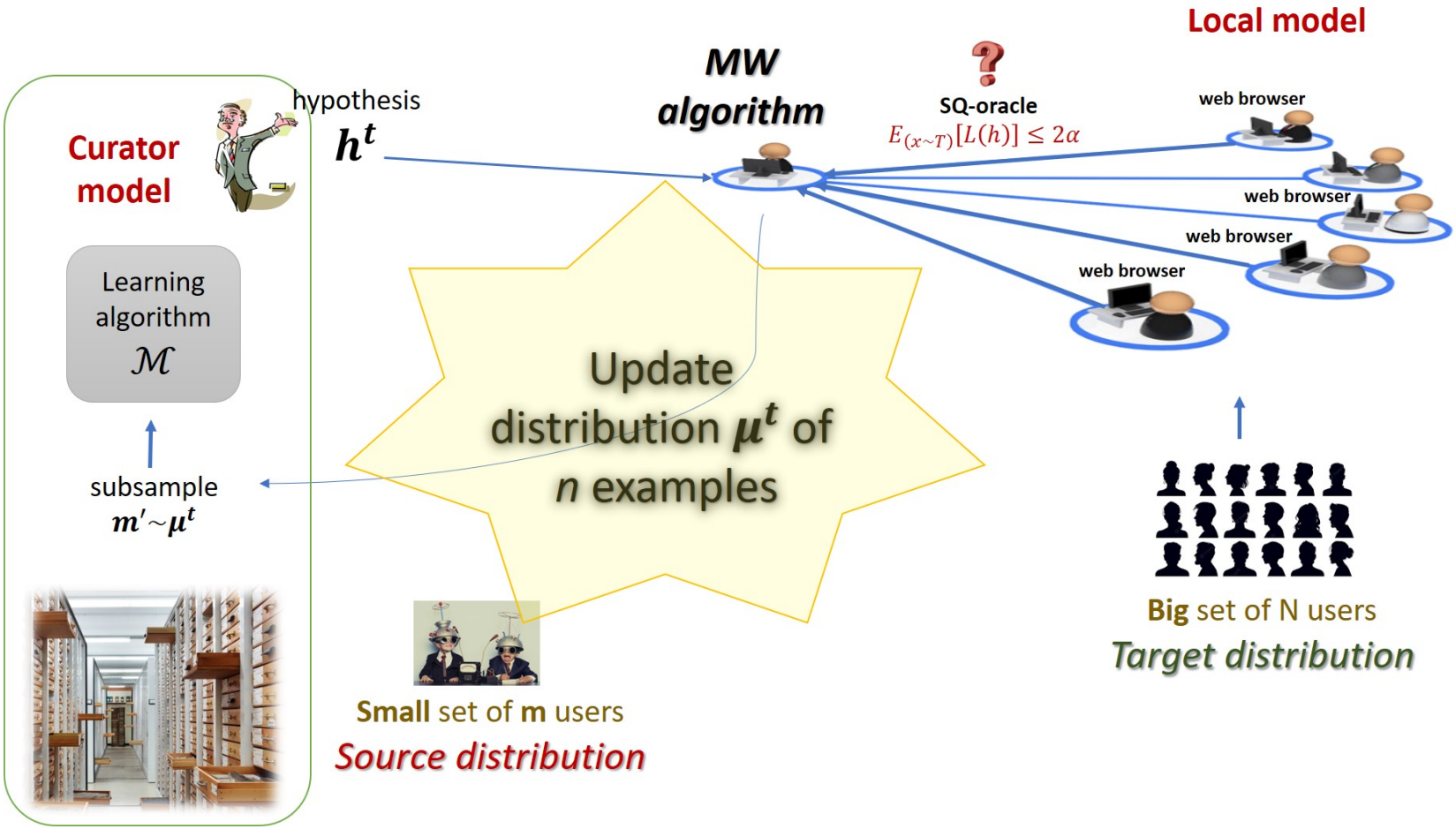


Subsample-Test-Reweigh - Private Multiplicative-Weights (MW) Algorithm for Transfer Learning



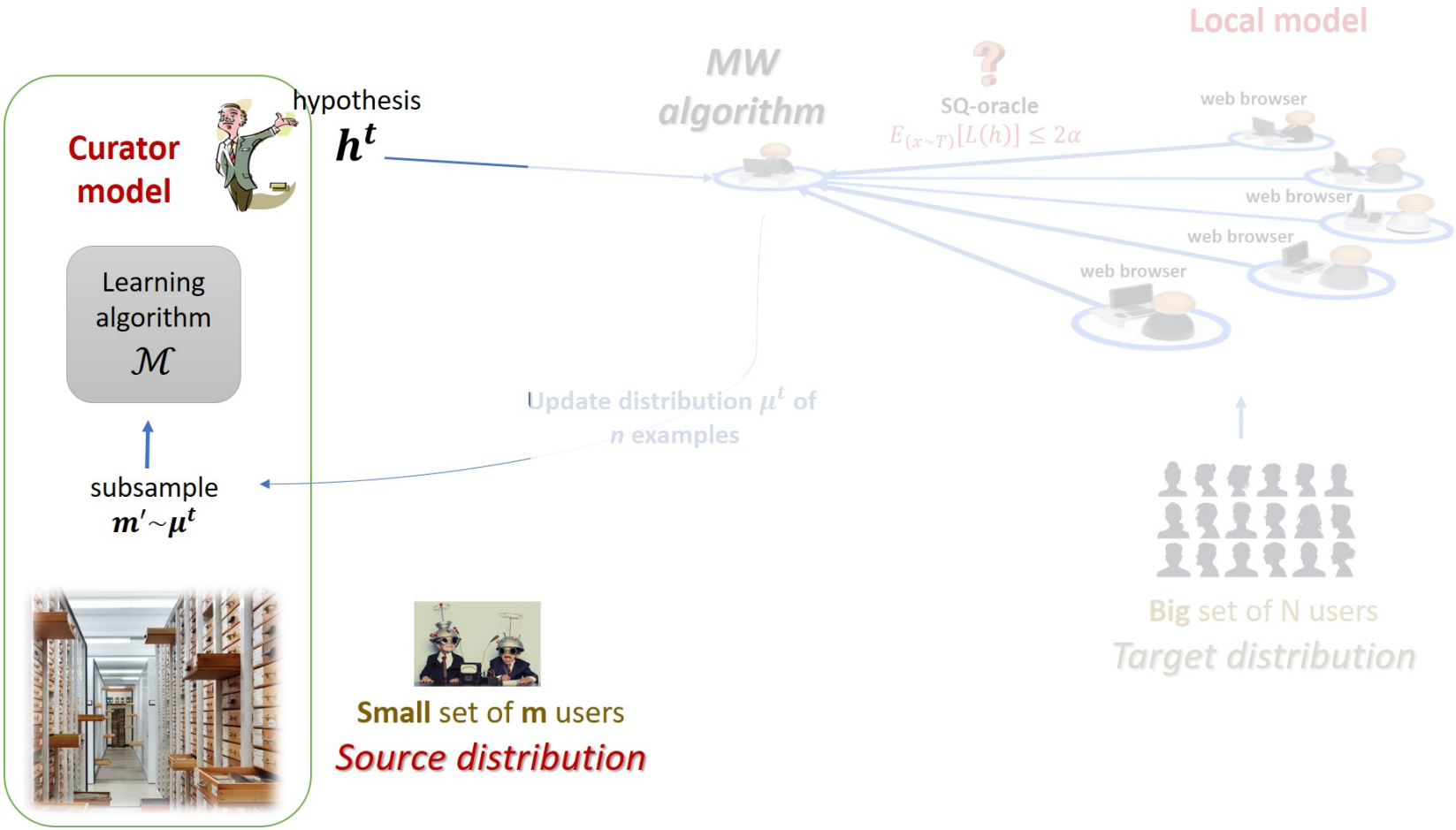
Mainly based on: Arora et al., 2012, Bun et al., 2020, 2018, Karwa & Vadhan, 2018, Cortes et al., 2010

Subsample-Test-Reweigh - Private Multiplicative-Weights (MW) Algorithm for Transfer Learning - step 1



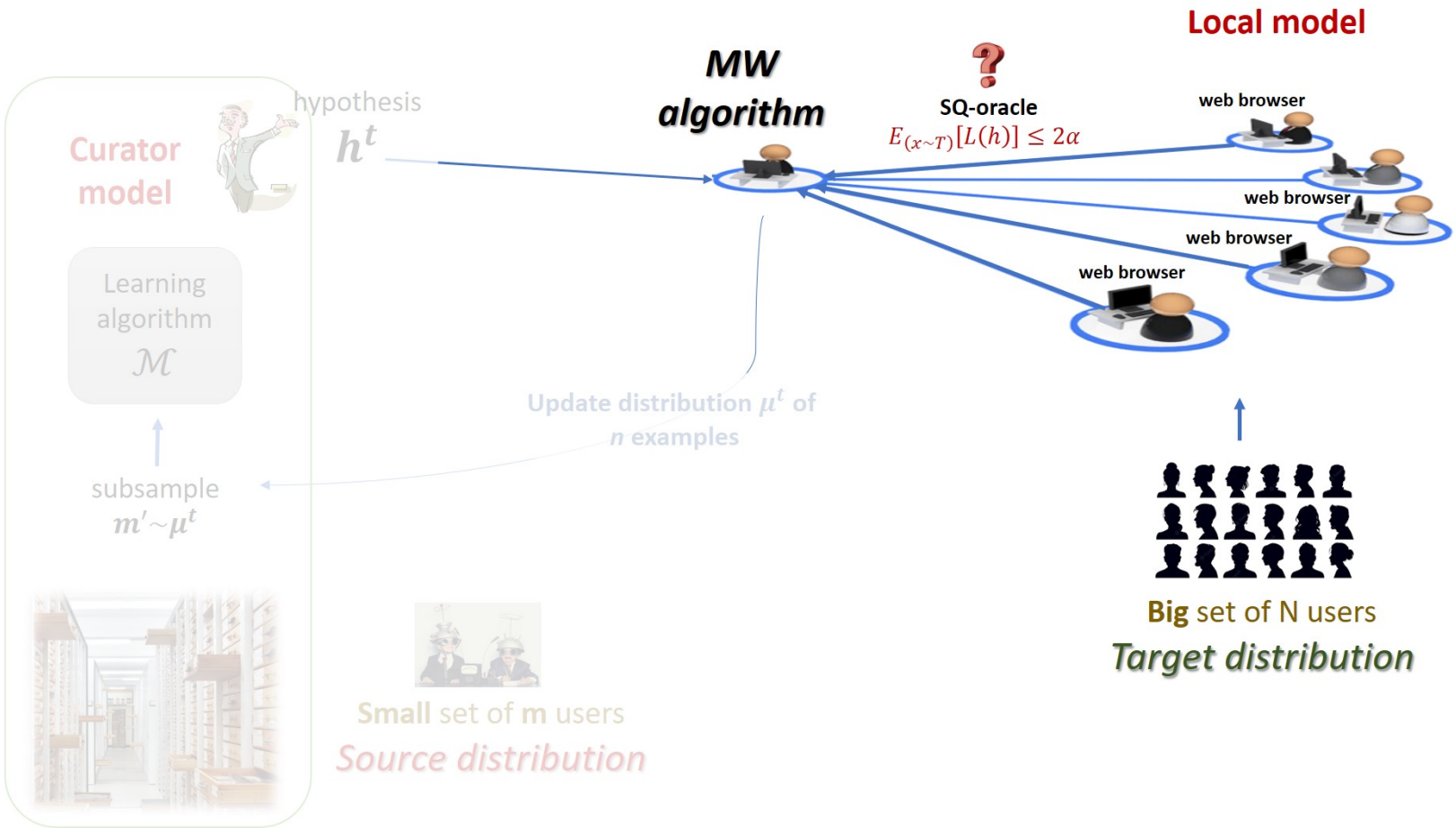
In each iteration MW creates a distribution μ^t on the examples from S....

Subsample-Test-Reweigh - Private Multiplicative-Weights (MW) Algorithm for Transfer Learning - step 2



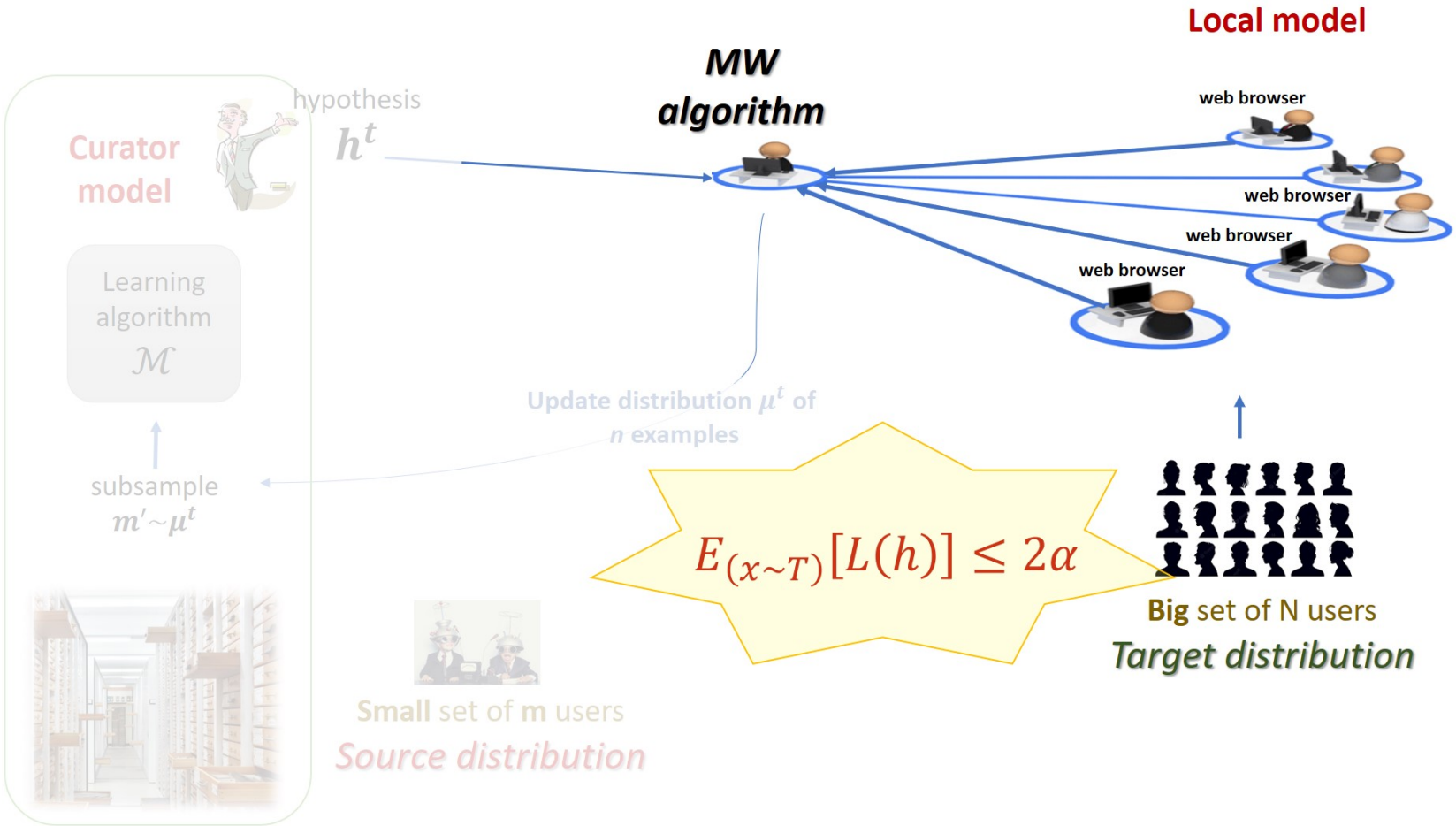
Runs a learning algorithm on a subsample from this μ^t distribution ...

Subsample-Test-Reweigh - Private Multiplicative-Weights (MW) Algorithm for Transfer Learning - step 3



Estimates the expected loss of the hypothesis h^t on the T distribution
If yes $\rightarrow$ halt and return h^t, if not $\rightarrow$ update the distribution and run again ...

Subsample-Test-Reweigh - Private Multiplicative-Weights (MW) Algorithm for Transfer Learning



The process is converges

Thanks For Listening!

Poster: Session 1 Track 1 - Hall E #914

Paper: https://arxiv.org/abs/2201.12018

Code: https://github.com/refael-kohen/SampleTestReweigh