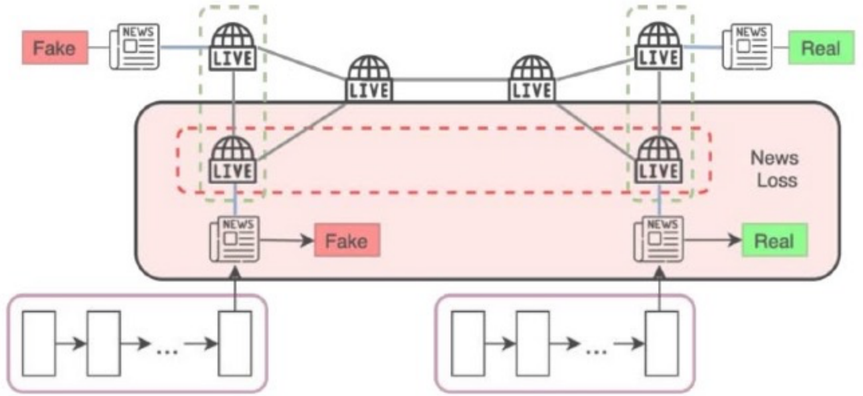
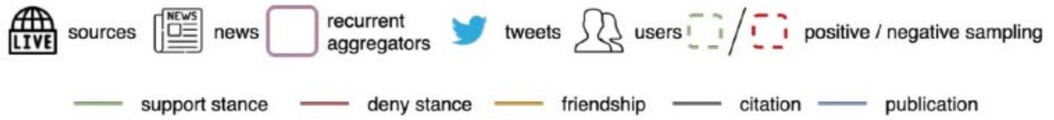
# Methodology

## FANG - Supervised Fake News Loss

- Combine the representation of an article and its source: $v_a = (z_a, z_s)$
- Passed through a fully connected layer: $o_a = \mathrm{W} v_a + b$
- Cross-entropy loss:
  - $\mathscr{L}_{\text{news}} = \frac{1}{T} \sum_a \left\{ y_a \cdot \log\left(\sigma\left(o_a\right)\right) + \left(1 - y_a\right) \cdot \log\left(1 - \sigma\left(o_a\right)\right) \right\}$



# Methodology

## FANG: Total loss function

- This is achieved by optimizing three concurrent losses:
  - Unsupervised Proximity Loss
  - Self-supervised Stance Loss
  - Supervised Fake News Detection Loss
- Define the total loss by linearly combining these three component losses:
- $\mathscr{L}_{\text{total}} = \mathscr{L}_{\text{prox}} + \mathscr{L}_{\text{stance}} + \mathscr{L}_{\text{news}}$