Use dictionaries on graphs for sparse classification/regression. E.g.: N documents in $\mathbb{R}^D$, compute multiscale dictionary Φ $(D \times M)$ on the D words. If f maps documents to their topic, write $f = X\Phi\beta + \eta$ and find β by

$$\mathrm{argmin}_\beta ||f - X\Phi\beta||_2^2 + \lambda||\{2^{-j\gamma}\beta_{j,k}\}||_1 \,,$$

which is a form of sparse regression. (λ, γ) are determined by cross-validation.