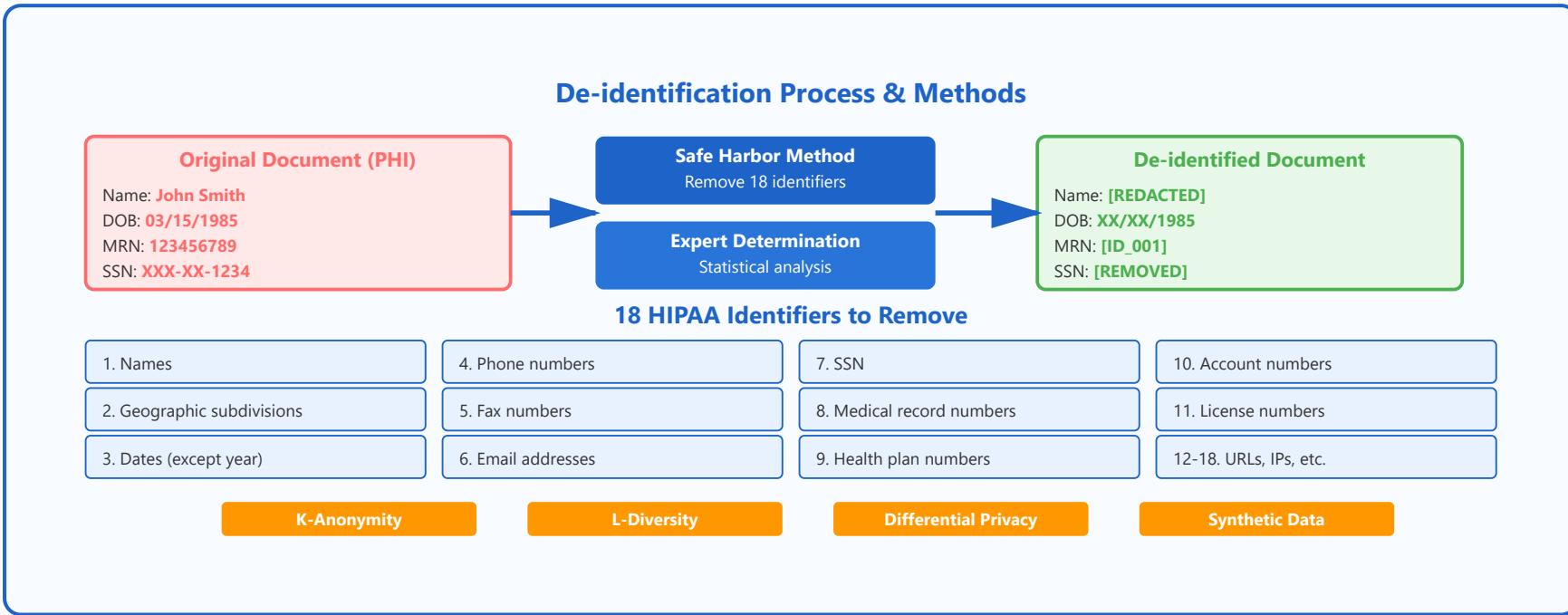


De-identification



Safe Harbor Method

- Remove all 18 identifiers
- Dates → year only
- Ages >89 → grouped as 90+
- Geographic: first 3 digits ZIP
- No statistical expertise needed



Expert Determination

- Statistical risk assessment
- Very small re-ID risk
- Retains more data utility
- Requires certified expert
- Document methodology



Automated Tools

- NER for PHI detection
- Philter, Scrubber, BoB
- Date shifting algorithms
- Validation required
- Hybrid human-AI review



Privacy Techniques

- K-anonymity (grouping)
- L-diversity (variety)
- Differential privacy (noise)
- Synthetic data (GANs)
- Utility vs privacy tradeoff