Two kinds of sensitive data:

1. data that identifies an individual

2. data that contains sensitive information about an individual or can be abused

Name is of type 1,

Health records are of type 2

Credit card numbers are of types 1 and 2

Plan:

1: build named entity recogniser on the following kinds of data: name, passport number, identification card number, bank account numbers, credit card numbers, loyalty program numbers, social security numbers

...customer numbers, usernames,...

special case: IP addresses & online behaviour

2: build text classifier as being on: finance, education, private, criminal, health, race, none of these

3: integrate: ID and one of these topics → highlight in output

more sophisticated: estimate number of distinct persons concerned & what kind of information is available on each → summarise

Notes: addresses only text & ID data, and not: Internet connectivity, geolocation, phone connectivity (i.e. the machine generated data), and biometric data

Algorithms:

1: presumably a recurrent neural network at character level (alternative for names: ngram vectorisation) →classification for (all numbers are just examples) 40 character substrings that overlap 30 characters with the previous substring if they contain identifying data

2: ?

3:?