Questo documento serve per testare **Docx Anonymizer**: il dizionario scansiona il documento cercando di individuare i nominativi, il testo viene linearizzato: il contenuto di formattazione (ad. Es tabelle) è gestito dal programma. Segue un testo con un po’ di nominativi:

Sono presenti questi **nominativi**:

1. Giancarlo Sibari
2. Filiberto De Luca
3. Carlo Pluto
4. Federico De Nicola
5. Francesco Paolo
6. Francesco Antonio
7. Antonio Franco
8. Antonio Andrea Luca
9. Lorenzo Mario Amorosa
10. Carmine de Rosa
11. Niccolò Dell’Uva

**Testo**:

Giancarlo Sibari, duca di Alessandria, decise di far parte di un testo senza senso. Con il suo amico Filiberto De Luca, tenta di fare un agguato al re Pluto Carlo. Attenzione! De Nicola Federico è un altro nominativo, e Francesco Paolo un altro ancora. Vi siete chiesti se il cognome sia “Francesco” o “Paolo”, quale termine è il cognome? Per ora, non importa: l’importante è distinguerlo da Francesco Antonio, o Antonio Franco. Sono tutte persone distinte. Anche Antonio Andrea Luca.

Attenzione alle omonimie: Antonio Luca non viene volutamente minimizzato se l’input lo specifica l’utente: per via della gestione delle omonimie!

Per apprezzare i falsi positivi individuati dal dizionario: Al Giovane è un nominativo...

Inoltre ecco come si comporta il programma con: Lorenzo Mario, Lorenzo Amorosa, Lorenzo Mario Amorosa, e solo il cognome Amorosa.

Altri due nominativi (notare minuscola, accento, apostrofo):

Carmine de Rosa, Niccolò Dell’Uva , Dell'UVA Niccolò e Dell’ Uva Niccolò.

E in una versione formattata: Carmine de Rosa, Niccolò Dell'Uva.

Infine una serie di nominativi molto simili tra loro per stressare il programma:Lorenzo Amorosa, Luca Amorosa, Lorenzo Mario Amorosa, Luca Mario, Mario Lorenzo.

(Nota: nei test-case sono stati espressi in input solo i nominativi indicati nella lista puntata precedente, alcuni dei nominativi presenti in quest’ultima riga saranno individuati solo con l’uso di dizionari)