

Normalizing Data

COmanage Registry supports the concept of *data normalization*. For example, if a petitioner enters the text " los angeles " into a field, normalization could correct that to "Los Angeles".

As of v0.9.4, normalization is enabled by default. Normalization is enabled on a per-CO basis, via *CO >> Configuration >> CO Settings*. Once enabled, normalization currently happens transparently, without meaningful user feedback (CO-282). After a supported form is submitted, normalizations are run before data is written to the database.

Currently, the following normalizations are supported:

Normalization	Description
mixCase	Make the first character in each word upper case and the remaining characters lower case. ⚠ For multibyte (non-ASCII7) support, PHP must be built with <code>--enable-mbstring</code> .
punctuationToSpace	Replace all non-alphanumeric characters with whitespace.
trimWhitespace	Leading and trailing whitespace is removed.

⚠ All defined normalizations are always run on the relevant data (models), when normalization is enabled for the CO.

Model / Field	mixCase	punctuationToSpace	trimWhitespace
Address.line1	✓		✓
Address.line2	✓		✓
Address.locality	✓		✓
Address.state	✓ (if string is at least 4 characters)		✓
Address.postal_code			✓
Address.country	✓ (if string is at least 4 characters)		✓
CoPersonRole.title	✓		✓
CoPersonRole.o	✓		✓
CoPersonRole.ou	✓		✓
CoPersonRole extended attributes of type VARCHAR			✓
EmailAddress.mail			✓
Identifier.identifier			✓
Name.honorific			✓
Name.given			✓
Name.middle			✓
Name.family			✓
Name.suffix			✓
TelephoneNumber.country_code		✓	✓
TelephoneNumber.area_code		✓	✓
TelephoneNumber.number		✓	✓
TelephoneNumber.extension		✓	✓
Url.url			✓ as of v3.1.0

Additional normalizations can be added by [writing a custom plugin](#). When adding new normalizations, note that `NormalizationBehavior` should run before `C hangelog Behavior`.