



Description of seqenv internal pipeline structure

Lucas Sinclair, lucas.sinclair@me.com

Last modified:

Tue Mar 01 2016

All numbers are fake
This document just describes a plan for processing data.

Dictionaries in the pipeline once run

text to counts		
"the deepest.."	-> (lake	-> 2.0)
"marine wat..."	-> (marine biome	-> 0.5,
	marine habit	-> 0.5,
	wetland	-> 1.0)
"My bath"	-> (

seq_to_counts	
lake	-> 2.0
marine biome	-> 0.5
marine habit	-> 0.5
wetland	-> 1.0

Result normalized	
This OTU is 50% lakish	
This OTU is 12.5% marine biomish	
This OTU is 12.5% marine habitish	
This OTU is 25% wetlandish	