

DATA GATHERING

Case Study on an issue resolved by the Connemara Programme

2016

Key Points

- Connemara data and information was diverse, dispersed and disassociated.
- The data and information existed in many formats the most important ones being hard copy, Online web sites, official databases and verbal.
- Priority was given to current information available digitally. Hard copy non current data and information was prioritized on its context importance for current data.
- New Photos had to be taken as the average age of existing stock was 5—8 years.

The Issue

Data and information on Connemara is held in many places and in many formats and The data and information on Connemara had never been systematically collected or codified for use in a database or presentation via digital means. Locating and normalizing the data and information was a major undertaking.

The Approach

The location, extraction, analysis and normalization of existing data and information is carried out in three main phases

1) Locate or Create: Where data or information was available in hard copy or soft copy it was gathered and placed in holder folders. Where resources (Photos, descriptions, geo location etc were required these were created by the Programme.

2) Data entry: The gross data and information was then scanned, or typed into a data base on which a word analysis script was use to auto generate top level categories.

3) Normalisation: Data sets were then normalized and aligned to ensure grammatical, content and presentation consistency which was critical for auto translation.

The Result

Over 4 million words and 80,000+ photos have been added, Normalised and categorised on the Programme's Single View Database. The data was located from hard copy (brochures, articles, advert, marketing etc) , web site, social media, open data sets, official government and EU sources and local Area subject experts.

Analysis of this (ongoing) effort has shown that the Programme has a 3 : 1 spend on data versus the technologies resources used to manage the data. No specialist skills were required for the data entry. Normalisation and Categorisation was done to fixed in-house rules