

METHOD FOR INFORMATION EXTRACTION FAILURE DIAGNOSIS AND AUTO-RECOVERY

5 **Abstract**

10 This invention relates to a computer-implemented method of refining rules for
extracting information comprising receiving data, retrieving a regular expression from an
extraction pattern database, annotating said regular expression, applying a first phase of
error detection extension having a plurality of first set of rules by subjecting said data to
said annotated regular expression and said plurality of first set of rules, generating a match
result from said application of said first phase of error detection extension, applying a
second phase of error detection extension having a plurality of second set of rules by
subjecting said data to said annotated regular expression and said plurality of second set of
rules, generating a mismatch result from said application of said second phase of error
15 detection extension in response to a mismatch in said data, applying a refinement process
to said data in response to generating said mismatch report from said second phase of
extraction of information from said data.

20 Figure 1 to accompany the abstract