ISO 24614-1:2010-11 (E)

Language resource management - Word segmentation of written texts - Part 1: Basic concepts and general principles

Contents

Page

Foreword iv

Introduction v

1 Scope 1

2 Terms and definitions 2

3 Basic framework for word segmentation 6

4 General principles of word segmentation 10

Annex A (informative) Representing word segmentation in XML 13

Bibliography 14