**Natural Language Processing for PDF/TIFF/Image Documents   
Computer Vision for Image Data  
SEGMENTATION MODULE  
High Precision Document Segmentation  
Technical Specification, Gap v0.9**

# 1 Segment

## 1.1 Segment Overview

The segment NLP preprocessor contains the following primary classes, and their relationships:

* X

**Columns**

**Segment (Base Class)**

**Paragraphs**

**Headings**

**Fig. 1a High Level view of Segment Class Object Relationships**

## 1.2 Segment Properties

The Segment class contains the following properties:

* TBD

## 1.3 Segment Overridden Operators

The following operators have their implementations (inherited from the base Object) class overridden:

* TBD

## 1.4 Segment Private Methods

The segment class contains the following private methods:

* TBD

## 1.5 Segment Public Methods

The Segment class contains of following public methods:

* There are no public methods.

## APPENDIX I: Updates

*Pre-Gap (Epipog) v1.4*

* An initial prototype was built

# APPENDIX II: Anticipated Engineering

The following has been identified as enhancement/issues to be addressed in subsequent update:

1. Add support for splitting dual column pages.
2. Add support for splitting paragraphs.
3. Add support for splitting headings.

**Proprietary and Confidential Information  
Copyright ©2018, Epipog, All Rights Reserved**