JANUS - A system for annotating, vetting and testing Natural Language Processing Systems

Michael Tanenblatt, Anni Coden, Igor Sominsky
IBM, T.J. Watson Research Center
Creation of annotated corpus
Visualization by concept

endoscopic biopsy (AE46-395890; 12/10/03): Invasive, grade 3 (of 4) adenocarcinoma identified.

Adenomatous polyp (AE46-39589H; 12/10/03): Tubulovillous adenoma with low-grade dysplasia, completely excised.

Resection specimen (AE46-32959; 12/23/03): Invasive, grade 3 (of 4) adenocarcinoma forms a fungating, circumscribed, ulcerated lesion involving muscularis propria. The regional and 2 separately submitted inferior mesenteric artery lymph nodes are negative for tumor. Surgical margins are positive for tumor.

Diagnoses: adenocarcinoma

Size: 7 x 3 x 3 cm

Primary Sites: rectum, sigmoid

Grade Values: 3 (of 4)
Specification of comparison

- <?xml version="1.0" encoding="UTF-8"?>
    - nullValueImage="" xsi:schemaLocation="http://www.apache.org/uima/cfe/config CFECfg.xsd ">
    - <tns:comparisonDescriptor>
      - <tns:truth className="truth" enclosingAnnotation="uima.tt.sentenceAnnotation">
        - <tns:targetAnnotationMatcher annotationTypeName="SCRAnatomicalSite"/>
        - <tns:featureAnnotationMatchers annotationTypeName="SCRAnatomicalSite" windowsizeInside="1">
          - <tns:groupFeatureMatchers>
            - <tns:featureMatchers refId="L1" featurePath="begin"/>
            - <tns:featureMatchers refId="L2" featurePath="end"/>
            - <tns:featureMatchers refId="L3" featurePath="code"/>
            - <tns:featureMatchers refId="L4" featurePath="coveredText"/>
          - </tns:groupFeatureMatchers>
        - </tns:featureAnnotationMatchers>
      - </tns:truth>
    - </tns:comparisonDescriptor>
    - </tns:CFEConfig>
Visualization of comparison

GROSS DESCRIPTION

FINAL DIAGNOSIS
Sigmoid, mass at 22 cm, endoscopic biopsy (AE46-39589O; 12/10/03): Invasive, grade 3 (of 4) adenocarcinoma.
Sigmoid, polyp, polypectomy (AE46-39589H; 12/10/03): Tubulovillous adenoma with low-grade dysplasia.
Sigmoid and rectum, resection specimen (AE46-32959; 12/23/03): Invasive, grade 3 (of 4) adenocarcinoma.
Multiple lymph nodes (16 regional and 2 separately submitted inferior mesenteric artery lymph nodes involved.)
Visual evaluation