The main goal of this thesis is the study of the actual state of eXtensible Hypertext Markup Language (for short XHTML) code errors in the world wide web Internet based on the violations of the recomendations made by The World Wide Web Consorcium (for short W3C). Not following these recomendations leads to problems while processing the document using standardizet tools. The second goal of this thesis is to create a software for converting the not-valid XHTML code into an as-valid-as-possible equivalent making it simpler to process.