

地方公共団体オープンデータ推進ガイドラインの データ作成例として構造化 PDF が追加

【2015年11月9日】

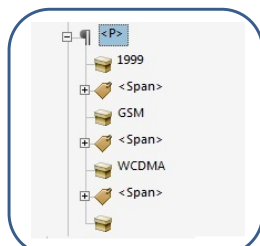
アドビ システムズ 株式会社（本社：東京都品川区、代表取締役社長：佐分利 ユージン、以下 アドビ）は、この度、内閣IT総合戦略本部主催の電子行政オープンデータ実務者会議により策定された「オープンデータをはじめよう～地方公共団体のための最初の手引書～」のオープンデータ作成例として構造化PDF[※]が追加されたことを発表しました。

行政機関の文書にはPDFが広く活用されていますが、PDFは一般に画像データとして認識されることが多く、検索やコンテンツの再利用が難しいという誤解からオープンデータの作成例に含まれていませんでした。しかし、PDF作成ツールによっては、データ構造化された機械判読可能なPDFも簡単に作成できることから、本年8月のガイドライン改訂において構造化PDFのデータ作成例に追加されました。

Adobe Acrobat DCは、電子文書PDFの国際規格であるISO32000-1に準拠した構造化PDFが作成できるため、オープンデータ対応に最適なソフトウェアです。

構造化PDFは「タグ」がつけられており、タグが示す文書の論理構造を利用することで内容の再利用性や検索性が格段に向上します。ワープロソフトで作った段組みや見出し、段落などの文書構造をPDFにも引き継いでいます。段組みがあっても見出しや段落の並び順が正確に認識できます。

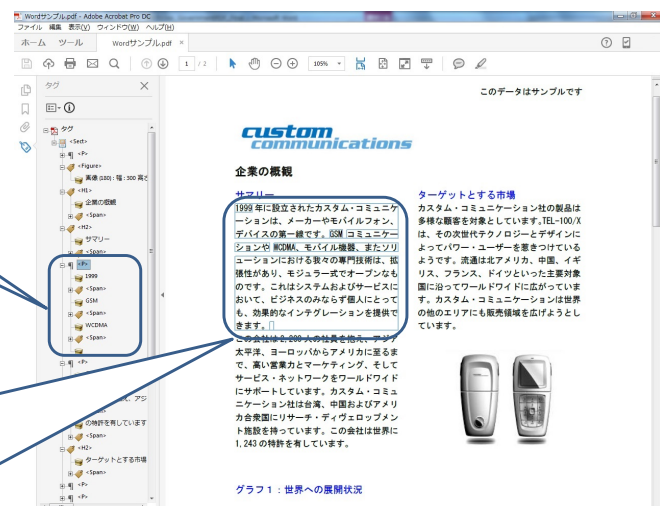
PDFの「タグ」構造



PDF本文の構造を解説

1999年に設立されたカスタム・コミュニケーションは、メーカーやモバイルフォン、デバイスの第一線です。GSM コミュニケーションや WCDMA、モバイル機器、またソリューションにおける我々の専門技術は、拡張性があり、モジュラー式でオープンなものです。これはシステムおよびサービスにおいて、ビジネスのみならず個人にとっても、効果的なインテグレーションを提供できます。

構造化PDFのサンプル(例)



アドビは、今回のガイドラインの改訂を機に、Adobe Acrobat DCによる構造化PDF作成の認知を拡大し、政府・自治体のオープンデータ推進を支援してまいります。

アドビ システムズ 株式会社 代表取締役社長 佐分利 ユージンは、「PDFは多くの行政機関で広く活用されている文書フォーマットです。このたび、オープンデータの作成例としてテキストの検索や抽出ができる構造化PDFが追加されたことで、市民や企業は自治体が所有する情報を効率的に閲覧・再利用することが可能になります。構造化PDFが簡単に作成できるツールであるAcrobat DCに関する認知を高める活動を通じて、自治体におけるデータ公開の推進を支援していきたいと考えております」と述べています。

<参考資料> オープンデータ推進ガイドラインにおけるPDFに関する記載

「オープンデータをはじめよう～地方公共団体のための最初の手引書～」（内閣官房IT総合戦略室作成）（P.31-32）データの作成 PDFの項目

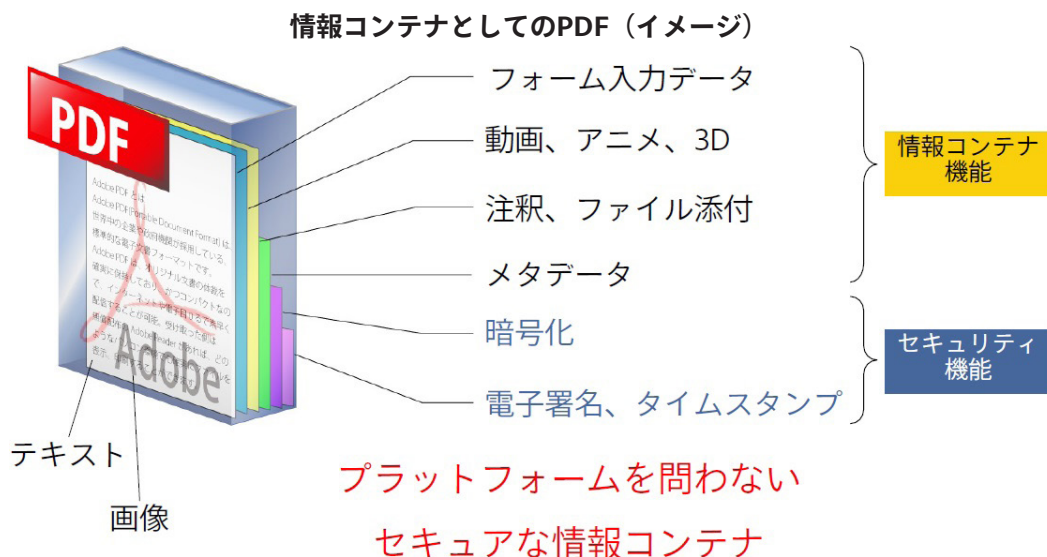
http://www.kantei.go.jp/jp/singi/it2/densi/kettei/opendata_tebikisyo.pdf

- ・ 「人が読んだり、印刷したりすることを念頭に置いてPDFファイルでデータを公開する場合でも、テキストの検索や抽出ができるPDFファイルにすることが望ましいです。」
- ・ 「文書構造などの情報が保持され、データの二次利用がしやすい「タグ付きPDF」を作る機能があるPDF作成ツールもあります。」

※ 構造化PDF

PDF文書の内容に、章、見出し、段落、図や表などといった文書の論理構造を示すタグを埋め込んだもので、タグ付きPDFとも呼ばれます。PDFは、もともとテキストデータ保持でき、テキストの抜き出しや検索ができるものですが、構造化PDFにすることにより、文書の構造や内容の繋がりが正確に表現できるため、機械判読性が高まり、複雑なレイアウトの文書でも高い精度でテキストの抜き出しや読み上げが行えるようになります。

なお、Acrobatを使えば、構造化PDFの作成だけでなく、他のアプリケーションで作成したファイルをPDF内に埋め込んだり、レイヤーを追加してテキストや画像を貼り付けたりできるなど、PDFを情報コンテナとして活用することができます。



アドビについて

アドビは、世界を動かすデジタル体験を提供します。

アドビ システムズ 株式会社はその日本法人です。同社に関する詳細な情報は、[Webサイト](#)に掲載されています。

© 2015 Adobe Systems Incorporated. All rights reserved. Adobe and the Adobe logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries. All other trademarks are the property of their respective owners.