

デジタルドキュメント(8)

高久雅生

2013年6月6日(木)3・4時限

(前回の復習 = ふりかえり)

- ドキュメントフォーマットとは？
- ドキュメントフォーマットの切り口
 - 書式・レイアウト, 複合コンテンツ, 標準化
- プレインテキストとその文字コード
- フォーマットの識別や判別
- ドキュメントフォーマットの例
 - メールフォーマット

前回の出席カード(質問)

- (プレインテキスト・文字コードの解釈について)
 - プレインテキストの読解が興味深かった。解答は「Hello World!」でよいのか？
 - Yes.
 - 文字コード関連で、文字化けについても少し伺いたかった(Unicodeではどんな言語も文字化けしないのか、など)
 - 文字化けは基本的に符号化方式(エンコーディング)の解釈の違いでおきます。
 - Unicodeであっても、正しい符号化方式を選択できない場合は、エラーとなります。
- (前回までのレポート評価、講評関連について)
 - A- はほぼ B と判断してよいか？
 - 理論的には、全課題においてA-である場合、最終評価は「B」となるはずです。
 - 課題について、レポートの要件は、「提供形態、具体例を説明せよ」との記述がされているのに、今回のスライドで先生がご提示された校正ミスの指摘や、詳細すぎる内容の説明(各章の組み立てのされ方、それぞれの内容の説明など)をするということは無駄であるように感じました。それは論文の要約であり、課題内容から論文の詳細を記述せよ、というような内容は汲み取りにくかったように思います。またA4一枚という制限から、論文によっては、要約ばかりの記述になり肝心となる形式等の部分が内容の薄いものになってしまうと思います。
 - デジタルドキュメントの説明として肝心の部分は文書の構成です。そして文書の構成は、文書内における図表や参考文献といった要素にどのようなものが含まれるか、どのような参照関係があるか、そしてデジタルメディア上でそれらがどのように表現されているかを説明することを指すと考えています。
 - 今さらレポートの書き方を教えてもらっても納得いかない。メディア創成でレポートといたらプログラムのソースコードと実行結果のこと。
 - 次回にいかせるように講評を加えていますので、評価に納得がいけないところがあれば担当者まで。追伸：自然言語でのレポートの書き方は重要です。

前回の出席カード(感想)

1. メールのフォーマットなど見慣れているけれど、改めて形式ばっていると思った
2. プレインテキストの読解が難しかった
3. レポートは3回くらいが良いです...
4. 2学期制に伴って、以前の3学期制よりもコマ数が多くなり、レポートが集中してしまい厳しいです。少数のレポート課題で評価していただけると助かります。4回以上となると、少し多すぎてしまうかなと...。弱音を吐いてしまいすみません...
5. 出席に関しては、学生は3分の2以上の出席をしない場合、単位はこないと既に分かっているはずなので、出席回数を満たしていない人に関しての処置をしても問題ないと思います
6. ドキュメントフォーマットというものについてよく理解できたと思う
7. 他の先生からの質問によって、私の中では、電子書籍といえば論文といった印象がついていることに気付いた。マンガ等を読むといった場面が想像できず、まだ私の中ではブームは来ないと思う
8. メールもデジタルドキュメントの一つだとは考えもしなかった
9. 今回の話は興味のある分野だったので面白かったです
10. テキストファイルの文字コードを判別しやすくするための工夫を知ることができてよかった
11. 日本の電子書籍は高いと思う

本日のお品書き

- デジタルドキュメントとドキュメントフォーマット
- (半)構造化文書
- HTML
 - 文書例
 - 要素、タグ、属性
 - ハイパーリンク
 - 構造と見栄え
 - 歴史と標準化
- XML
 - 文書例
 - 整形式

ドキュメントフォーマット (2)

(再掲)ドキュメントフォーマットの切り口 (1)

- テキスト (text) vs バイナリー (binary)
 - ビットデータ
 - 文字コードによる解釈
 - 外字
- フォーマットの指定・識別・判別
- シンプルコンテンツ vs 複合メディア
 - 埋め込みコンテンツ
 - ハイパーリンク
- メタデータ
 - 埋め込みメタデータ
 - 外部メタデータ記述
- 文書レイアウト
 - ページ概念
- 文書内の書式要素
 - 見栄え / スタイル
 - フォント
- ファイル vs ストリーム
 - データの保存・蓄積と配信

(再掲)ドキュメントフォーマットの切り口 (2)

- オープンフォーマット
 - 移植可能性 / ソフトウェア独立性
 - Free / proprietary
- 標準化
 - デファクト標準とデジュール標準 (“de facto” vs “de jure”)
- 文書フォーマットのバージョン
- フォーマット変換
 - テキスト → HTML
 - LaTeX → PDF
- 圧縮
 - 可逆 vs 非可逆
- セキュリティ
 - パスワード
 - 電子署名
- 長期保存
- デジタルフォレンジック

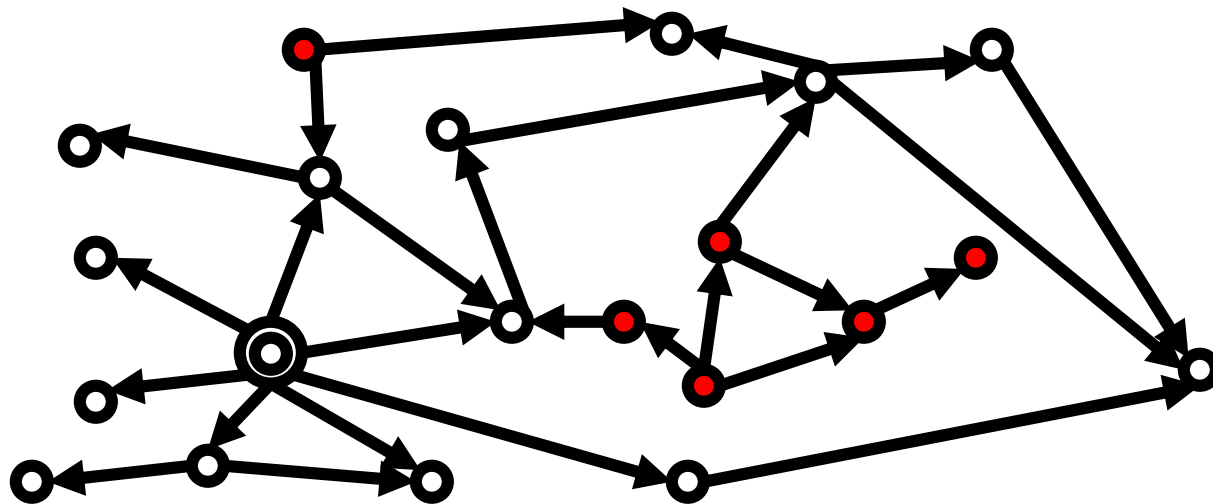
HTML (Hypertext Markup Language)

- ウェブ上でのコンテンツ記述用言語
- テキストデータのやり取り
- ウェブブラウザによる解釈と描画
- タグ; メタデータ; 構造
- ハイパーメディア; ハイパーテキスト
– リンク
– ページ埋め込み
- マルチメディア (画像、音声、動画)
- ダイナミックな表現; インタラクション

```
<html>
<head>
<title>タイトル</title>
</head>
<body>
<h1>見出し</h1>
<p>段落</p>
</body>
</html>
```

WebとHTML

- HTTP, URI, HTMLの3点によるWebの実現！
- HTTPレイヤでのデータ転送とドキュメントフォーマット指定
- URI指定によるリンクを通じたハイパーメディアの実現



HTMLの歴史

- HTMLの発展
 - HTML 1.1 (1992) World Wide Webの黎明
 - HTML 2.0 (1995) 標準化 (IETF: RFC 1866)
 - HTML 3.2 (1997) W3Cによる標準化 (デファクトの追認)
 - HTML 4 (1997) 整理統合、国際化
- XHTML (2000)
 - XMLに基づく再出発
- HTML5 (2014?)
 - 成熟 (混沌?) からの新しい出発

オープンシステムとしてのウェブ / HTML テキストフォーマットとしてのウェブ / HTML

- テキストベースのフォーマットであるので、自由にソフトウェア (Webブラウザ) を作成し、提供できる。

– 多様なブラウザの例

↓ Google Chrome

w3m →



テキストによる文書例

デジタルドキュメント(2013年・春)

2013年度春学期に開講する「デジタルドキュメント」科目についての情報を提供します。

授業予定

注意: 進捗状況に応じて、適宜変更します。

1. 第1回: デジタルドキュメントとは? [第1回レポート課題]
2. 第2回: 学術研究分野におけるデジタルドキュメント(その1)
3. 第3回: 学術研究分野におけるデジタルドキュメント(その2)
[第2回レポート課題]
4. 第4回: 学術研究分野におけるデジタルドキュメント(その3)
5. 第5回: 電子書籍とその利用環境(その1)
6. 第6回: 電子書籍とその利用環境(その2) [第3回レポート課題]
- ...

HTMLフォーマットによる文書例

<h1>デジタルドキュメント(2013年・春)</h1>

<p>2013年度春学期に開講する「デジタルドキュメント」科目についての情報を提供します。</p>

<h2>授業予定</h2>

<p>注意: 進捗状況に応じて、適宜変更します。</p>

第1回: デジタルドキュメントとは? [第1回レポート課題]

第2回: 学術研究分野におけるデジタルドキュメント(その1)

第3回: 学術研究分野におけるデジタルドキュメント(その2) [第2回レポート課題]

第4回: 学術研究分野におけるデジタルドキュメント(その3)

第5回: 電子書籍とその利用環境(その1)

第6回: 電子書籍とその利用環境(その2) [第3回レポート課題]

HTMLフォーマットの表示例

The screenshot shows a web browser window with the address bar displaying `masao.jpn.org/lecture/2013/digital-document/`. The page has a light green background and a blue dashed border around the main content area. The browser's address bar includes navigation buttons (back, forward, home, search) and a search engine icon. Below the address bar, there is a bookmark bar with various links like '<ref>タグ生成', 'MPDL', 'delicious', 'Import to Mendeley', '科学情報室ホーム...', 'Mapping Statistics', 'ブックマークレット', and 'その他のブックマーク'. The page content starts with a breadcrumb trail: `HOME > 講義資料 > デジタルドキュメント(2013年・春)`. To the right, it shows the publication date: '初版公開日: 2013年04月15日' and the last update date: '最終更新日: 2013年06月05日'. The main title is 'デジタルドキュメント(2013年・春)'. Below the title, a paragraph states: '2013年度春学期に開講する「デジタルドキュメント」科目についての情報を提供します。'. A section titled '授業予定' (Lecture Schedule) follows, with a note: '注意: 進捗状況に応じて、適宜変更します。'. The schedule lists 10 lectures, each with a date, topic, and a link to '授業資料' (Lecture Material). The topics are: 1. Digital Document and its format (Lecture 1 report topic), 2. Academic research fields and digital documents (Lecture 2 report topic), 3. Academic research fields and digital documents (Lecture 3 report topic), 4. Academic research fields and digital documents (Lecture 4 report topic), 5. Electronic books and their usage environment (Lecture 5 report topic), 6. Electronic books and their usage environment (Lecture 6 report topic), 7. Digital document format (Lecture 7 report topic), 8. Digital document format (Lecture 8 report topic), 9. Digital document format (Lecture 9 report topic), 10. Digital document format (Lecture 10 report topic). A section titled 'シラバス記載事項' (Syllabus Items) follows. At the bottom, there is a table with two columns: '授業科目名' (Lecture Subject Name) and '科目' (Subject). The first row shows 'デジタルドキュメント' (Digital Document) under '授業科目名' and '科目'.

masao.jpn.org/lecture/2013/digital-document/

<ref>タグ生成 MPDL delicious Import to Mendeley 科学情報室ホーム... Mapping Statistics ブックマークレット その他のブックマーク

HOME > 講義資料 > デジタルドキュメント(2013年・春)

初版公開日: 2013年04月15日
最終更新日: 2013年06月05日

デジタルドキュメント(2013年・春)

2013年度春学期に開講する「デジタルドキュメント」科目についての情報を提供します。

授業予定

注意: 進捗状況に応じて、適宜変更します。

- 第1回(4/18): デジタルドキュメントとは? (授業資料) [第1回レポート課題]
- 第2回(4/25): 学術研究分野におけるデジタルドキュメント(その1) (授業資料)
- 第3回(5/2): 学術研究分野におけるデジタルドキュメント(その2) (授業資料) [第2回レポート課題]
- 第4回(5/9): 学術研究分野におけるデジタルドキュメント(その3) (授業資料)
- 第5回(5/16): 電子書籍とその利用環境(その1) (授業資料)
- 第6回(5/23): 電子書籍とその利用環境(その2) (授業資料) [第3回レポート課題]
- 第7回(5/30): デジタルドキュメントのフォーマット(その1) (授業資料)
- 第8回(6/6): デジタルドキュメントのフォーマット(その2)
- 第9回(6/13): デジタルドキュメントの総合芸術: つながるコンテンツとメタデータ
- 第10回(6/20): デジタルドキュメントの将来像

シラバス記載事項

授業科目名	デジタルドキュメント
科目	

HTMLにおけるタグと要素

- ブラケット文字: < と > に囲んだ中にタグ名
- h1タグ → <h1>
- 開始タグと終了タグ
 - <h1> ~ </h1>
- 要素
 - 開始タグと終了タグで囲まれた部位を「要素」と呼ぶ
 - 入れ子構造をとることができる

第1回:デジタルドキュメントとは？[第1回レポート課題]



HTMLにおけるハイパーリンク

```
<a href="http://www.tsukuba.ac.jp/">筑波大学の公式サイト</a>
```

ブラウザでの表示→

[筑波大学の公式サイト](http://www.tsukuba.ac.jp/)

- a (Anchor)タグによるURL参照を通じたリンク
 - href属性の値としてURLを記述してリンクを表現
- a要素の内容はアンカーテキストと呼ばれるリンク部分となる
 - リンク先との関係を示す記述がなされる
 - → Webサーチエンジンにおける利用



http://commons.wikimedia.org/wiki/File:US_Navy_Anchor_Insignia.svg

HTMLにおけるハイパーリンク(フレーム)

```
<html>
<frameset rows="172,*">
  <frame src="title.html" name="ue">
  <frameset cols="150,*">
    <frame src="sidemenu.html" name="hidari">
    <frame src="welcome.html" name="migi">
  </frameset>
</frameset>
</html>
```

- Framesetタグにより、フレーム枠全体を設定し、frameタグにより、実際の各フレームにおいて参照するページURLを指定
 - 例: <http://www.slis.tsukuba.ac.jp/mediaman/>
- 利点:
 - 単一ページ単位をシンプルに保て、ナビゲーション用として同一内容を保持して示したりできる。
- 欠点:
 - 各フレームにおける個別ページが孤立してしまいがち。個別ページに外部からリンクを付与したりすることが難しい。

HTMLにおけるハイパーリンク(画像)

- 画像:img要素

```
<a href="dl.html"></a>
```

```
<a href="http://klis.tsukuba.ac.jp/assets/files/Syllabus20130401.pdf"></a>
```

```
<a href="http://www.tsukuba.ac.jp/" title="筑波大学"></a>
```

- alt, title属性による
テキスト情報の追加

<http://www.klis.tsukuba.ac.jp/>



HTMLにおけるハイパーリンク (複合オブジェクト)

- その他: 動画や外部ページの埋め込み等
- `iframe`: 任意のコンテンツをその場に埋め込む

```
<iframe width="800" height="200" src="http://klis.tsukuba.ac.jp"></iframe>
```

— (Transclude)



授業予定

注意: 進捗状況に応じて、適宜変更します。

saveMLAK Screencast1

SCREENCASTOMATIC

Start Recording

Screen Recorder

Help | Send Feedback

Ready!


For best results, set the volume on your microphone before recording (Windows/Mac only).

You can close this page or navigate to another site if you need to.

SSDメーカー

良い品質 低価格

1. 第1回 (4/18): デジタルドキュメントとは? (授業資料) [第1回レポート課題]
2. 第2回 (4/25): 学術研究分野におけるデジタルドキュメント(その1) (授業資料)
3. 第3回 (5/2): 学術研究分野におけるデジタルドキュメント(その2) (授業資料) [第2回レポート課題]
4. 第4回 (5/9): 学術研究分野におけるデジタルドキュメント(その3) (授業資料)
5. 第5回 (5/16): 電子書籍とその利用環境(その1) (授業資料)
6. 第6回 (5/23): 電子書籍とその利用環境(その2) (授業資料) [第3回レポート課題]



localhost/~masao/test/z3.html

HOME > 講義資料 > デジタルドキュメント(2013年・春)

初版公開日: 2013年04月15日
最終更新日: 2013年06月05日

デジタルドキュメント(2013年・春)

2013年度春学期に開講する「デジタルドキュメント」科目についての情報を提供します。

授業予定

注意: 進捗状況に応じて、適宜変更します。

筑波大学
University of Tsukuba

知識情報・図書館学類 情報学群
College of Knowledge and Library Sciences, School of Informatics

1. 第1回 (4/18): デジタルドキュメントとは? (授業資料) [第1回レポート課題]
2. 第2回 (4/25): 学術研究分野におけるデジタルドキュメント(その1) (授業資料)
3. 第3回 (5/2): 学術研究分野におけるデジタルドキュメント(その2) (授業資料) [第2回レポート課題]
4. 第4回 (5/9): 学術研究分野におけるデジタルドキュメント(その3) (授業資料)
5. 第5回 (5/16): 電子書籍とその利用環境(その1) (授業資料)
6. 第6回 (5/23): 電子書籍とその利用環境(その2) (授業資料) [第3回レポート課題]

(余談)ハイパーメディア/ハイパーテキストの夢

WWWの勃興, 普及

- ハイパーメディア概念の発見と普及へ
 - Memex (Vannevar Bush) - 1945
 - Xanadu (Ted Nelson) - 1963?
 - WWW (Tim Berners-Lee) – 1989
- WWWが捨てたもの!?
 - 閲覧・編集の統合(ウィキ?)
 - バージョン管理
 - 多様で拡張可能なリンク
 - 著作権管理



ティム・バーナーズ=リー:「Webの創成: World Wide Webはいかにして生まれどこに向かうのか」. 高橋徹訳. 毎日コミュニケーションズ, 2001, 279p. (ISBN: 4-8399-0287-9)

HTMLの構造と見栄え

- 文書構造, 文書内の要素
 - 見出し
 - 段落
 - リスト
 - 表
 - 画像
 - リンク
- コンテンツ内容と見栄えの分離
 - 論理スタイル要素と物理スタイル要素
 - `強調` vs `強調`
 - `強調` vs `<i>強調</i>`
 - ``要素の廃止... `強調`
→ オープンフォーマットとの親和性

スタイルシートの例 (CSS)

- 文字書式やスタイルを指定した変更

```
body {  
  background-color: #cfc;  
  margin: 1% 2% 0px;  
  line-height: 1.2;  
}  
h2 {  
  padding-top: 0.1em;  
  padding-bottom: 0.1em;  
  border-style: solid;  
  text-indent: 0.5em;  
  color: #223322;  
  background-color: #eeeeee;  
  border-color: #669966;  
  border-width: 0 0 thin thick;  
}
```

The screenshot shows a web browser window with the URL `masao.jp/lecture/2013/digital-document/`. The page title is "デジタルドキュメント(2013年・春)". Below the title, there is a section titled "授業予定" (Lecture Schedule) which is highlighted with a dashed blue border. The schedule lists 10 lectures with their dates and topics, each followed by a link to "授業資料" (Lecture Materials). Below the schedule, there is a section titled "シラバス記載事項" (Syllabus Information). At the bottom, there is a table with details about the course.

授業科目名	デジタルドキュメント
科目番号	GE70301
単位数	2.0 単位
標準履修年次	3・4 年次
時間割	春AB 木3.4

XML (Extensible Markup Language)

- 順序付き木構造による表現
- 整形式 (Well-formedness)
- スキーマ (文書内容の定義) に則った、個別のインスタンス言語
 - XHTML
 - その他の文書形式
- 参照関係
- 文字コード: Unicode
- 元情報としてのXML文書
 - 変換することにより、他のデータ形式へ
 - One source multiple use...



```

22 ↓
23 <manualpage metafile="getting-started.xml.meta">↓
24   <title>Getting Started</title>↓
25   <summary>↓
26   <p>If you're completely new to the Apache HTTP Server, or even to running↓
27   a website at all, you might not know where to start, or what questions to↓
28   ask. This document walks you through the basics.</p>↓
29 </summary>↓
30 ↓
31 <section id="clientserver">↓
32 <title>Clients, Servers, and URLs</title>↓
33 <p>↓
34 Addresses on the Web are expressed with URLs – Uniform Resource Locators↓
35 – which specify a protocol (e.g. <code>http</code>), a servername (e.g. ↓
36 <code>www.apache.org</code>), a URL-path (e.g. ↓
37 <code>/docs/current/getting-started.html</code>), and possibly a query↓
38 string (e.g. <code>?arg=value</code>) used to pass additional↓
39 arguments to the server. ↓
40 </p>↓
41 <p>A client (e.g., a web browser) connects to a server (e.g., your Apache
42 HTTP Server), ↓
43 with the specified protocol, and makes a <strong>request</strong> for a re
44 source using the↓
45 URL-path.</p>↓
46 ↓
47 <p>The URL-path may represent an

```

<http://httpd.apache.org/docs/trunk/getting-started.xml>

```
- <manualpage metafile="getting-started.xml.meta">
  <title>Getting Started</title>
  - <summary>
    <p>If you're completely new to the Apache HTTP Server, or even to running
      website at all, you might not know where to start, or what questions to
      ask. This document walks you through the basics.</p>
  </summary>
  - <section id="clientserver">
    <title>Clients, Servers, and URLs</title>
    - <p>
      Addresses on the Web are expressed with URLs - Uniform Resource
      Locators - which specify a protocol (e.g.
      <code>http</code>
      ), a servername (e.g.
      <code>www.apache.org</code>
      ), a URL-path (e.g.
      <code>/docs/current/getting-started.html</code>
      ), and possibly a query string (e.g.
      <code>?arg=value</code>
      ) used to pass additional arguments to the server.
    </p>
    - <p>
      A client (e.g., a web browser) connects to a server (e.g., your Apache
      HTTP Server), with the specified protocol, and makes a
      <strong>request</strong>
      for a resource using the
      http://httpd.apache.org/docs/trunk/getting-started.xml

```

← → C:\Users\masao\Desktop\m ... Getting Started - Apach... x C:\Users\masao\Desktop... 新しいタブ

Modules | Directives | FAQ | Glossary | Sitemap

Apache HTTP Server Version 2.5

Apache > HTTP Server > Documentation > Version 2.5

Getting Started

Available Languages: en

If you're completely new to the Apache HTTP Server, or even to running a website at all, you might not know where to start, or what questions to ask. This document walks you through the basics.

Clients, Servers, and URLs

Addresses on the Web are expressed with URLs - Uniform Resource Locators - which specify a protocol (e.g. `http`), a servername (e.g. `www.apache.org`), a URL-path (e.g. `/docs/current/getting-started.html`), and possibly a query string (e.g. `?arg=value`) used to pass additional arguments to the server.

A client (e.g., a web browser) connects to a server (e.g., your Apache HTTP Server), with the specified protocol, and makes a **request** for a resource using the URL-path.

The URL-path may represent any number of things on the server. It may be a file (like `getting-started.html`) a handler (like `server-status`) or some kind of program file (like `index.php`). We'll discuss this more below in the [Web Site Content](#) section.

The server will send a **response** consisting of a status code and, optionally, a response body. The status code indicates whether the request was successful, and, if not, what kind of error condition there was. This tells the client what it should do with the response. You can read about the possible response codes in [HTTP Server wiki](#).

Details of the transaction, and any error conditions, are written to log files. This is discussed in greater detail below in the [Logs Files and Troubleshooting](#) section.

Hostnames and DNS

- [Clients, Servers, and URLs](#)
- [Hostnames and DNS](#)
- [Configuration Files and Directives](#)
- [Web Site Content](#)
- [Log Files and Troubleshooting](#)
- [What Else Do I Need To Know?](#)
- [Comments](#)

<http://httpd.apache.org/docs/trunk/getting-started.xml>

まとめ

- デジタルドキュメントとドキュメントフォーマット
- (半)構造化文書
- HTML
 - 要素、タグ、属性
 - ハイパーリンク
 - 構造と見栄え
 - 歴史と標準化
- XML
 - メタ言語
 - 整形式
 - スキーマ

出席票 及び レポートの提出

提出年月日、学籍番号、所属、氏名、感想コメント等(あれば)を記入のうえ、提出してください。

提出位置:

3編生
2013xxxxx

3編生
2012xxxxx

3年次
(知識)
2011xxxxx

3年次
(創成)
2011xxxxx

4年次
2010xxxxx,
etc.

※前回以前に欠席し、返却を受けていないレポート課題がある者は申し出て、返却レポートを受けとること。