



*A Data Usability Company*  
**ANTENNA HOUSE**

# 組版結果の自動分析

白鯨（Moby-Dick）  
初版を再現

**Tony Graham**

XML Division

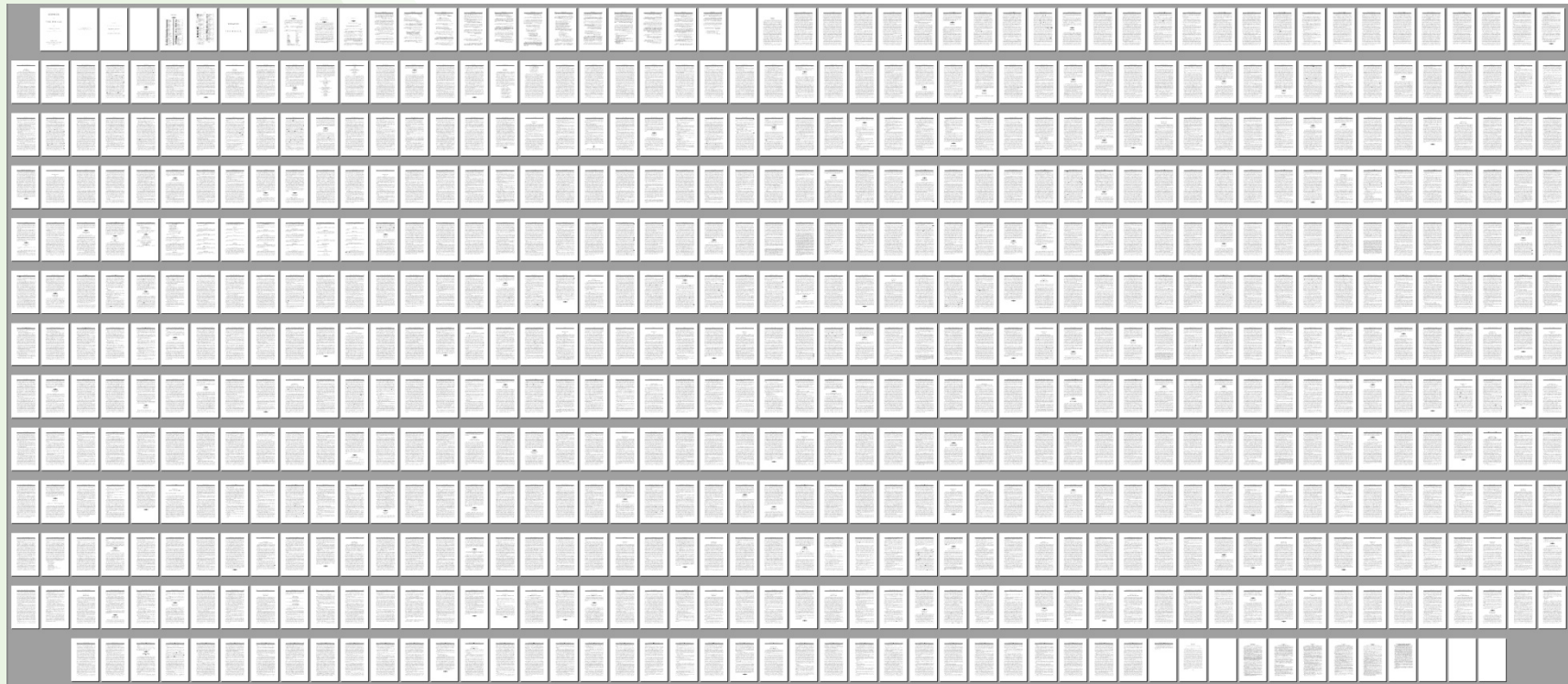
Antenna House, Inc.

tgraham@antenna.co.jp

tony@antennahouse.com

@tgraham\_antenna

# Moby-Dick; or, The Whale (白鯨)



# 組版結果の自動分析

- 何を、そしてなぜ
- 白鯨初版
- XMLソース
- 組版
- 組版結果

# 何を、そしてなぜ？

- 自動分析機能のテスト環境
  - 自動分析は手作業より速い
- 白鯨のアメリカ初版：
  - EPUBとCSSのサンプルに頻繁に使われる
  - 著作権が切れている
  - テキストはXMLとして利用可能
  - スキャンされたオリジナルのページが利用可能

# 人の手か、それとも機械？

妥当かどうかの判断が電子的に行われる場合、人間の想像力や判断力があれば回避できた不規則性が発生する可能性が高いです。鈍感な読者でもイライラするでしょう。したがって、機械による分析項目は人間による選択肢よりも制限されるべきでしょう。

Methods of Book Design

Hugh Williamson, 1983



# Antenna House Formatter による自動分析

- 巻末の空白ページ
- ハイフンで終わる連続する行
- 同じ単語が行頭や行末で連続する行
- ブロックの前後の行数
- ページウィドウ
- 段落ウィドウ
- リバー（川）
- 不均衡な見開きページ
- 空白

# 白鯨の用途は…

- 「ラテン文字のレイアウトとページ付けの要件」のサンプル
- EPUB のサンプル
- Lorem Ipsum の代替
- コーディング演習と課題

# 再現性を求めて

- Project Gutenberg
  - 多くの EPUB ソースがある
  - 詳細が不十分
- Melville Electronic Library
  - カラー ページ画像だが、TEI XML は公開されていない
- Wright American Fictionプロジェクト
  - TEI XML であり白黒のページ画像

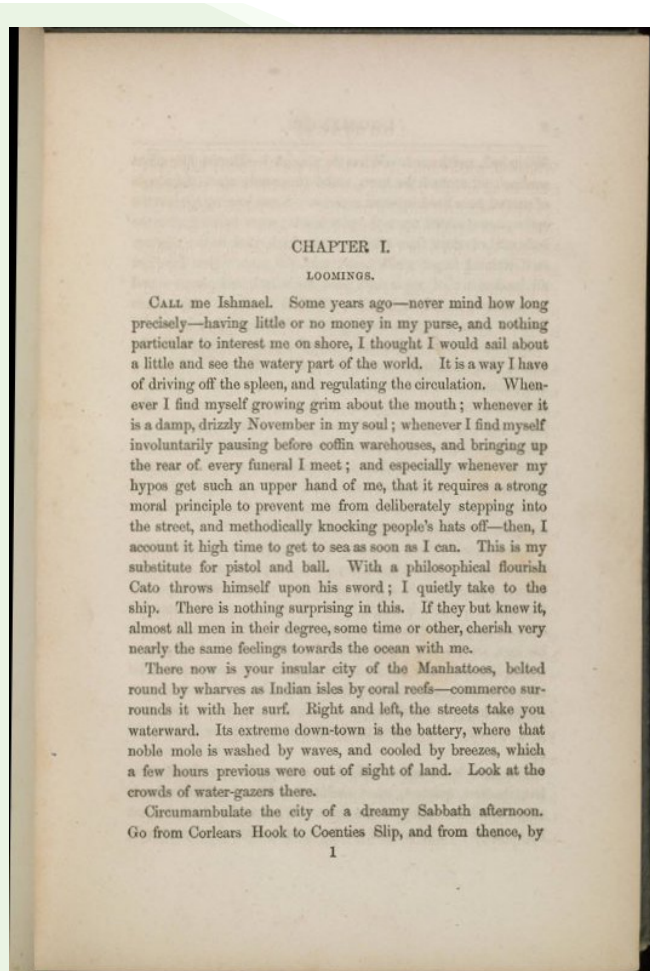


# Wright American Fiction

- 2000年に発案
- Indiana University（と関連する大学）
- 750,000ページ以上をスキャン
- OCR
- 約 3,000冊、1,200冊は完全にエンコード済み
- TEI Lite P3 SGML → TEI P4 XML → TEI P5 XML

# Moby-Dick (白鯨) TEI

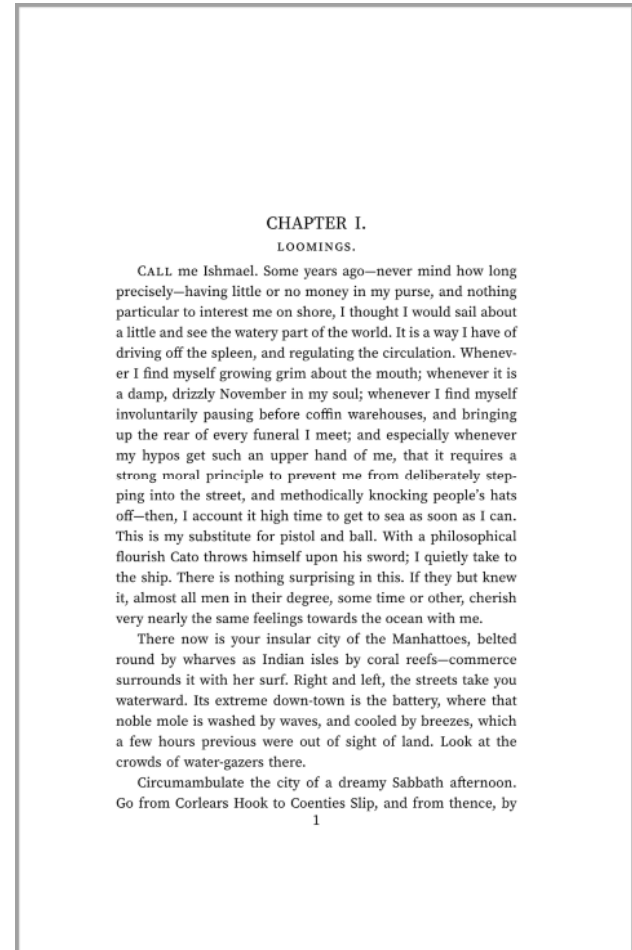
- 斜体テキストに空の `<hi rend="i"/>`
  - *Folio* のような外来語の単語は除く
- スモールキャピタルには大文字を使った
- ヘッダーに `<fw>` がない
- ページ末尾がハイフンで終わるページが複数
- 句読点が不正確。£ ではなく ¥ が使用された



**Pastiche n. (パスティーシュ 名詞.)**

他の作品、芸術家、または時代のスタイルを模倣した芸術作品。

**Pastiche v. (パスティーシュ 動詞)**  
芸術家または作品を模倣する。

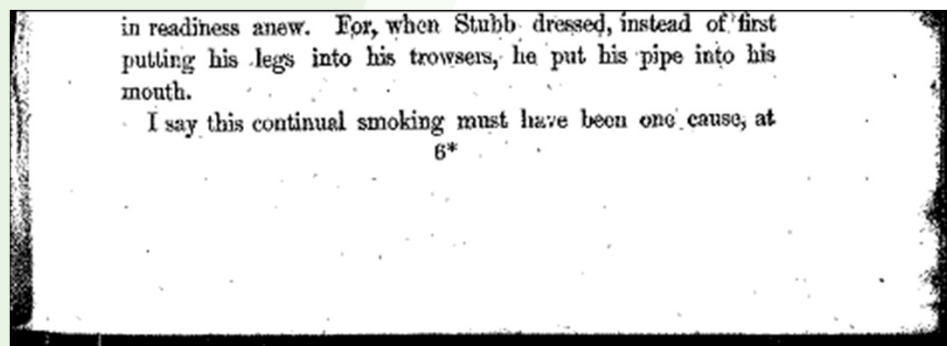


# どのくらい模倣できたか？

- 異なるフォントとフォントサイズ
- 異なる空白の規則

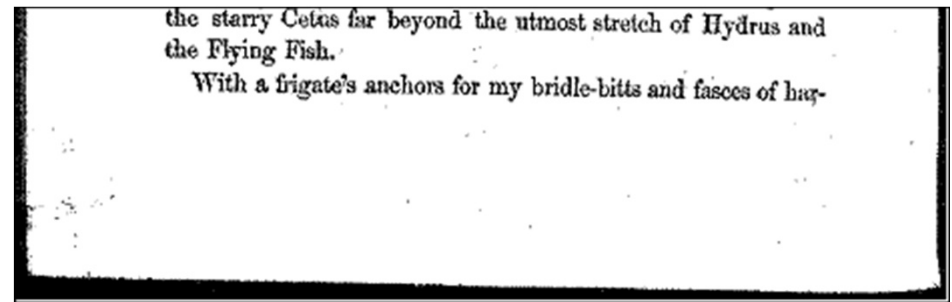
“SHIP, ahoy! Hast seen the White Whale?”

- 一行のオーファン

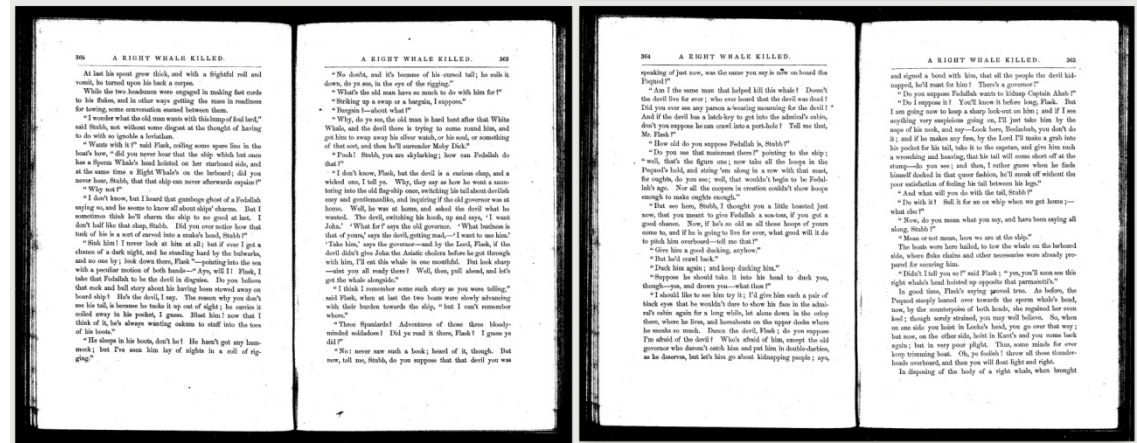


# どのくらい模倣できたか？

- ページ末尾のハイフン

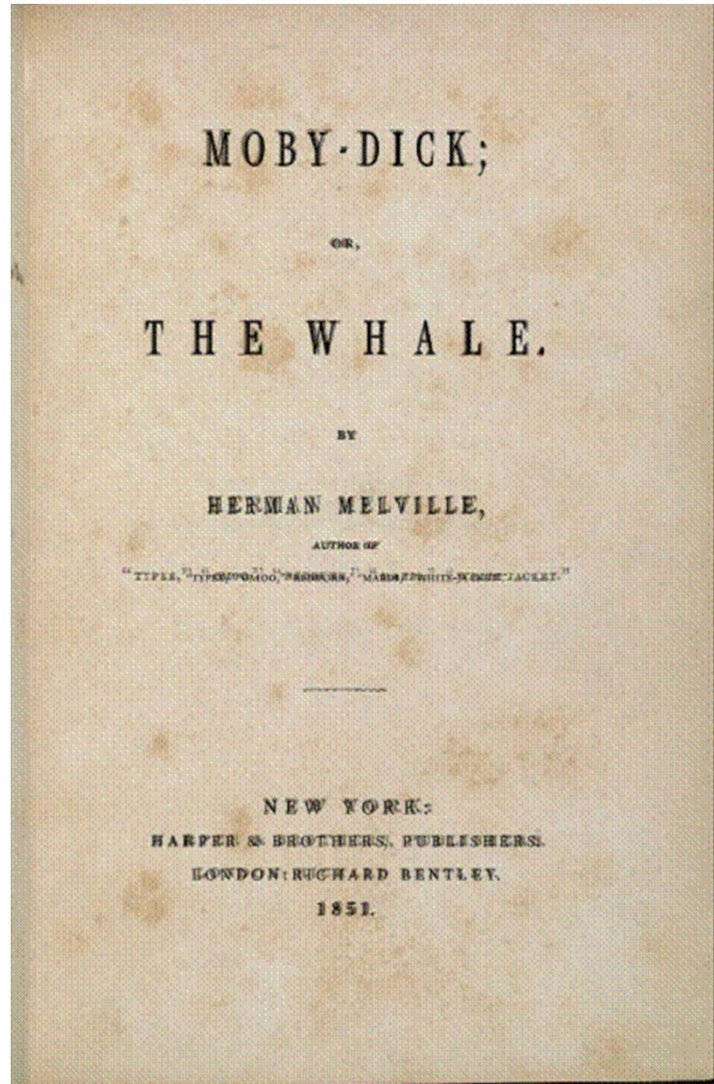


- ページが足りない



# 段階的に近似

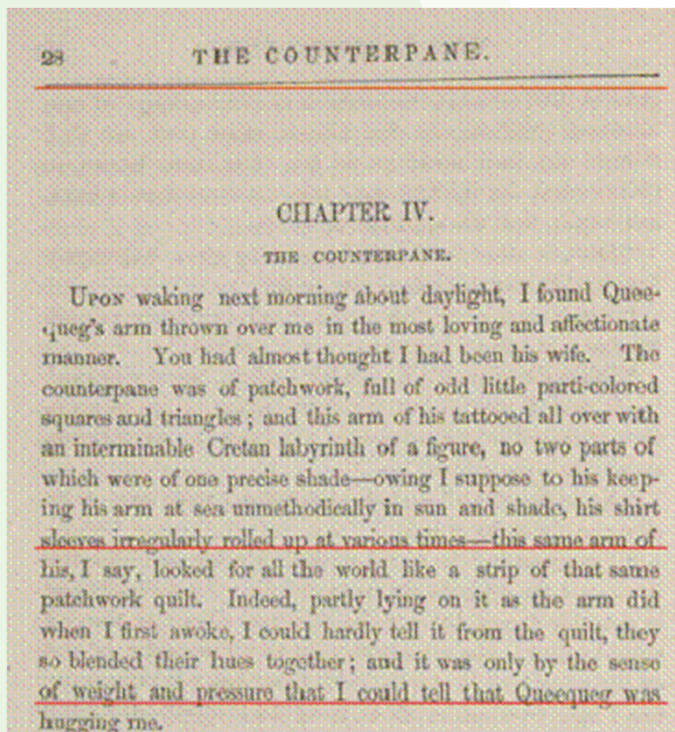
- 初稿：大まかな近似
- 変更を加えて確認する
  - 満足のいくまで繰り返す
  - いくつかは差し戻し
  - 設計でカバーできない文脈もある



Company  
HOUSE

# 背景としてのページ画像

ページ画像をまっすぐにする



背景として設定する

```
<fo:simple-page-master
  master-name="PageMaster"
  page-height="7.375in"
  page-width="4.875in"
  background-image=
    "page-images/MD_Amer_0038.jpg"
  axf:background-size="5.21in"
  background-position=
    "-0.12in -0.15in"
  axf:bleed="0.5in"
  axf:crop-offset="0.5in">
```



# 前付

- タイトル
- 著作権
- 献辞
- 目次
- フライト
- タイ
- 源
- 抜
- 粹

MOBY-DICK; OR, THE WHALE. BY HERMAN MELVILLE. NEW YORK: HARPER & BROTHERS, PUBLISHERS, LONDON, STRAND ROAD, 1851.								IN THREE VOLS. BY HIS NEPHEW, EDU WARD, IN REVISED EDITION. NATHANIEL HAWTHORNE.								CONTENTS.								MOBY-DICK; OR, THE WHALE.							
EXTRACTS								ETYMOLOGICAL								EXTRACTS								EXTRACTS							
EXTRACTS								EXTRACTS								EXTRACTS								EXTRACTS							
EXTRACTS								EXTRACTS								EXTRACTS								EXTRACTS							
EXTRACTS								EXTRACTS								EXTRACTS								EXTRACTS							



# タイトルとフライタイトル

```
<docTitle>
```

```
  <titlePart>MOBY-DICK;</titlePart>
```

```
  <titlePart type="sub">OR, THE  
  WHALE.</titlePart>
```

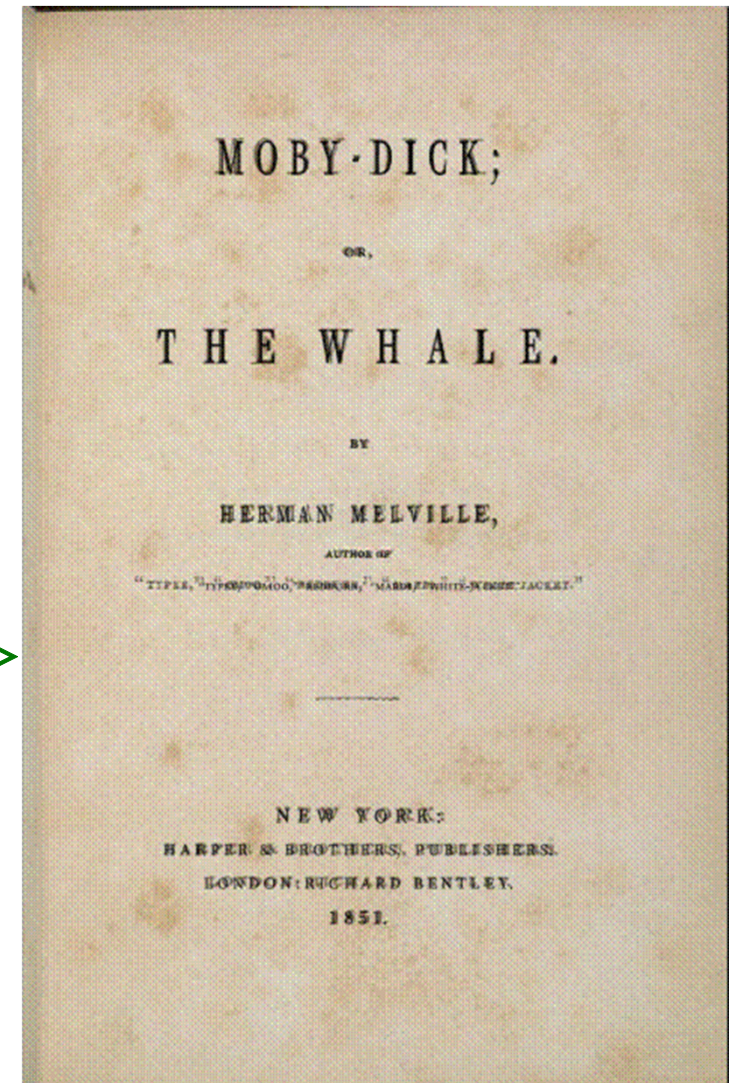
```
</docTitle>
```

```
<div type="fly_title">
```

```
  <head>MOBY-DICK; OR, THE WHALE.</head>
```

```
</div>
```

- マークアップを無視する
- `xsl:analyze-string` を  
 使用して FO を生成する



CONTENTS.

xciii.—The Castaway. . .	458
xciv.—A Squeeze of the Hand. . . . .	463
xcv.—The Cassock . . .	467
xcvi.—The Try-Works . .	468
xcvii.—The Lamp . . . .	474
xcviii.—Stowing Down & Clearing Up . . .	474
xcix.—The Doubloon. . .	478
c.—The Pequod meets the Samuel Enderby of London.	485
ci.—The Decanter . . .	493
xxx.—Queen Mab . . . . .	343
xxxi.—Cetology . . . . .	344
xxxii.—The Speckysunder . . .	359
xxxiii.—The Cabin Table . . .	362
xxxiv.—The Mast-Head . . . .	366
xxxv.—The Quarter-Deck. Akab and all . . . . .	376
xxxvi.—Gannet . . . . .	384
xxxvii.—Dark . . . . .	386
xxxix.—First Night-Watch . . .	387
xl.—The Crunch . . . . .	322
xli.—Stalk's Supper . . . . .	324
xlii.—The Whale as a Dish . . .	333
xliii.—The Shark Massacre . . .	336
xliv.—Cutting In . . . . .	338
xlv.—The Blanket . . . . .	340
xlvi.—The Funeral . . . . .	343
xlvii.—The Sphynx . . . . .	345

# 目次

```
<div type="contents">
  <pb n="v (Table of Contents) " xml:id="VAC7237-0003"/>
  <head>CONTENTS.</head>
  <list>
    <item>I.—Loomings. <ref target="VAC7237-0013"
      rend="right">1</ref>
    </item>
    <item>II.—The Carpet Bag. <ref target="VAC7237-0017"
      rend="right">7</ref>
    </item>
```

- <pb> への参照が一部不正確
- 相互参照が可能な目次

# 引用

“Silly Mansoul swallowed it without chewing, as if it had been a sprat in the mouth of a whale.”

*Pilgrim's Progress.*

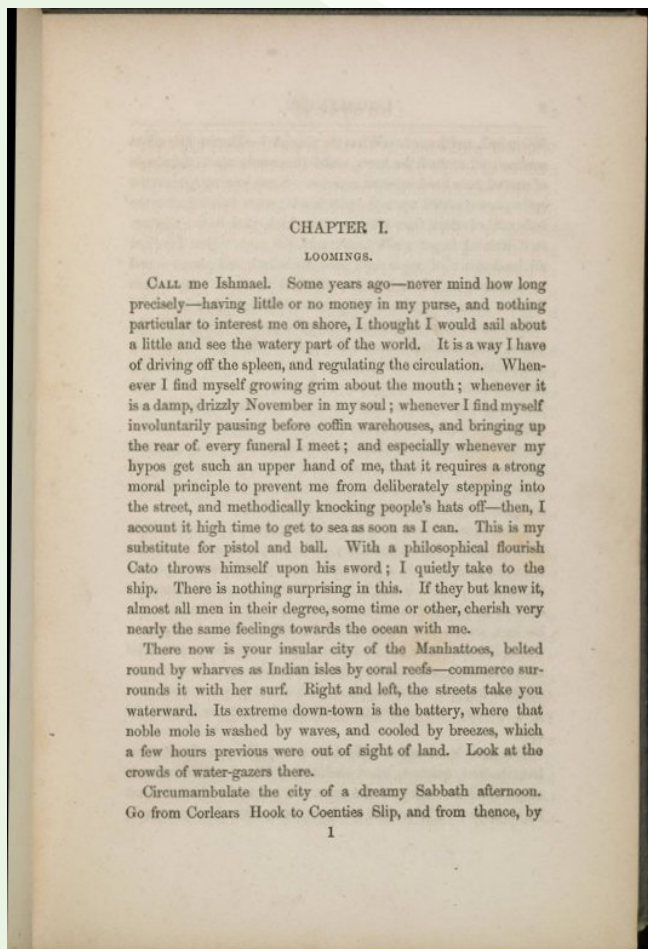
“Very like a whale.”

*Hamlet.*

```
<cit>
  <q>
    <p>"Very like a whale."</p>
  </q>
  <bibl>
    <title>Hamlet</title>.</bibl>
</cit>
```

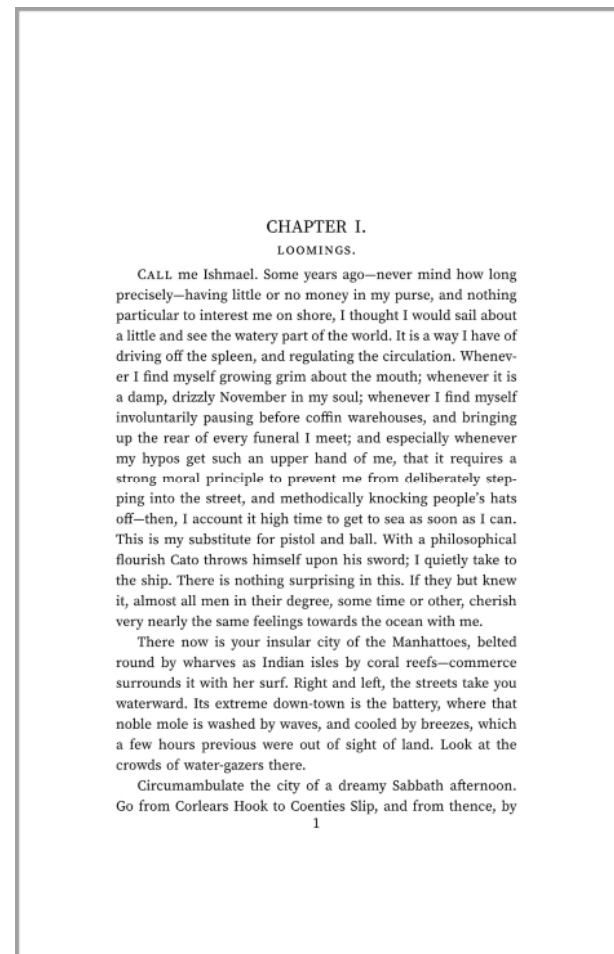
# <p>を使用して<bibl>の子を処理

```
<xsl:template match="q/p">
  <fo:block>
    <xsl:apply-templates />
    <xsl:if test="position() = last() and
                exists(..following-sibling::*[1][self::bibl])">
      <fo:leader leader-pattern="space"/>
      <fo:leader leader-pattern="space" leader-length.optimum="100%"/>
      <fo:inline-container padding-left="2em" padding-right="0.125in"
                          max-width="80%" text-indent="0">
        <fo:block text-align="right">
          <xsl:apply-templates
            select="..following-sibling::*[1]/node()" />
        </fo:block>
      </fo:inline-container>
    </xsl:if>
  </fo:block>
</xsl:template>
<xsl:template
  match="bibl[exists(preceding-sibling::*[1][self::q[p]])]"
  priority="5" />
```



# 本文

- 135 の章
- 画像がひとつ
- 碑銘
- 曲
- スピーチと舞台演出



# 章区切り

But already the sable wing was before the old man's eyes; the long hooked bill at his head: with a scream, the black hawk darted away with his prize.

An eagle flew thrice round Tarquin's head, removing his cap to replace it, and thereupon Tanaquil, his wife, declared that Tarquin would be king of Rome. But only by the replacing of the cap was that omen accounted good. Ahab's hat was never restored; the wild hawk flew on and on with it; far in advance of the prow: and at last disappeared, while from the point of that disappearance, a minute black spot was dimly discerned, falling from that vast height into the sea.



## CHAPTER CXXXI.

### THE PEQUOD MEETS THE DELIGHT.

THE intense Pequod sailed on; the rolling waves and days went by; the life-buoy-coffin still lightly swung; and another ship, most miserably misnamed the Delight, was descried. As she drew nigh, all eyes were fixed upon her broad beams, called shears, which, in some whaling-ships, cross the quarter-deck at the height of eight or nine feet; serving to carry the spare, unrigged, or disabled boats.

Upon the stranger's shears were beheld the shattered, white ribs, and some few splintered planks, of what had once been a whale-boat; but you now saw through this wreck, as plainly as you see through the peeled, half-unhinged, and bleaching skeleton of a horse.

"Hast seen the White Whale?"

"Look!" replied the hollow-cheeked captain from his taff-rail; and with his trumpet he pointed to the wreck.

"Hast killed him?"

"The harpoon is not yet forged that will ever do that," answered the other, sadly glancing upon a rounded hammock on the deck, whose gathered sides some noiseless sailors were busy in sewing together.

"Not forged!" and snatching Perth's levelled iron from the crotch, Ahab held it out, exclaiming—"Look ye, Nantucketer; here in this hand I hold his death! Tempered in blood, and tempered by lightning are these barbs; and I swear to temper them triply in that hot place behind the fin, where the White Whale most feels his accursed life!"

"Then God keep thee, old man—see'st thou that"—pointing to the hammock—"I bury but one of five stout men, who were alive only yesterday; but were dead ere night. Only that one I bury, the rest were buried before they died; you sail upon their tomb." Then turning to his crew—"Are ye ready there? place the plank then on the rail, and lift the body, so, then—Oh! God"—advancing towards the hammock with uplifted hands—"may the resurrection and the life—"

"Brace forward! Up helm!" cried Ahab like lightning to his men.

But the suddenly started Pequod was not quick enough to escape the sound of the splash that the corpse soon made as it struck the sea; not so quick, indeed, but that some of the flying bubbles might have sprinkled her hull with their ghostly baptism.

As Ahab now glided from the dejected Delight, the strange life-buoy hanging at the Pequod's stern came into conspicuous relief.

"Ha! yonder! look yonder, men!" cried a foreboding voice in her wake. "In vain, oh, ye strangers, ye fly our sad burial; ye but turn us your taffrail to show us your coffin!"



## CHAPTER CXXXII.

### THE SYMPHONY.

IT was a clear steel-blue day. The firmaments of air and sea were hardly separable in that all-pervading azure; only, the pensive air was transparently pure and soft, with a woman's look, and the robust and man-like sea heaved with long, strong, lingering swells, as Samson's chest in his sleep.

Hither, and thither, on high, glided the snow-white wings of small, unspckled birds; these were the gentle thoughts of the feminine air; but to and fro in the deeps, far down in the bottomless blue, rushed mighty leviathans, sword-fish, and sharks; and these were the strong, troubled, murderous thinkings of the masculine sea.

But though thus contrasting within, the contrast was only in shades and shadows without; those two seemed one; it was only the sex, as it were, that distinguished them.

Aloft, like a royal czar and king, the sun seemed giving this gentle air to this bold and rolling sea; even as bride to groom. And at the girthing line of the horizon, a soft and tremulous motion—most seen here at the equator—denoted the fond, throbbing trust, the loving alarms, with which the poor bride gave her bosom away.

Tied up and twisted; gnarled and knotted with wrinkles; haggardly firm and unyielding; his eyes glowing like coals, that still glow in the ashes of ruin; untottering Ahab stood forth in the clearness of the morn; lifting his splintered helmet of a brow to the fair girl's forehead of heaven.

Oh, immortal infancy, and innocence of the azure! Invisible winged creatures that frolic all round us! Sweet childhood of air and sky! how oblivious were ye of old Ahab's close-coiled

# 章区切りを自動生成

```
<xsl:template
  match="div[@type = 'chapter']
    [exists(head[@type = 'sub'] |
      fw[@type = 'head'])]">
  <xsl:if
    test="exists(preceding-sibling::div
      [@type = current()/@type])">
    <fo:block axf:suppress-if-first-on-page="true"
      space-after="0.2in"
      space-after.precedence="force">
      <fo:external-graphic src="separator.svg" />
    </fo:block>
  </xsl:if>
  <fo:block
    id="{@type}-{count(preceding::div
      [@type = current()/@type]) + 1}">
```

```
<fo:block
  id="{@type}-{count(preceding::div
    [@type = current()/@type]) + 1}">
  <fo:block-container
    axf:baseline-grid="none"
    axf:baseline-block-snap="none"
    keep-together.within-page="always"
    keep-with-next.within-page="always"
    space-before="{
      if (exists(preceding::div[1]
        [@type = 'chapter']))
        then '0.5in'
        else '0.72in'}"
    space-before.conditionality="retain">
    <xsl:apply-templates select="head" />
  </fo:block-container>
  <xsl:apply-templates select="* except head" />
</fo:block>
</xsl:template>
```



# クジラサイズの脚注

Iroquois, the midwinter sacrifice of the sacred White Dog was by far the holiest festival of their theology, that spotless, faithful creature being held the purest envoy they could send to the Great Spirit with the annual tidings of their own fidelity; and though directly from the Latin word for white, all Christian priests derive the name of one part of their sacred vesture, the alb or tunic, worn beneath the cassock, and though among the holy pomps of the Romish faith, white is specially employed in the celebration of the Passion of our Lord, though in the Vision of St. John, white robes are given to the redeemed, and the four-and-twenty elders stand clothed in white before the great white throne, and the Holy One that sitteth there white like wool, yet for all these accumulated associations, with whatever is sweet, and honorable, and sublime, there yet lurks an elusive something in the innermost idea of this hue, which strikes more of panic to the soul than that redness which affrights in blood.

This elusive quality it is, which causes the thought of whiteness, when divorced from more kindly associations, and coupled with any object terrible in itself, to heighten that terror to the furthest bounds. Witness the white bear of the poles, and the white shark of the tropics; what but their smooth, flaky whiteness makes them the transcendent horrors they are? That ghastly whiteness it is which imparts such an abhorrent mildness, even more loathsome than terrific, to the dumb gloating of their aspect. So that not the Serce-fanged tiger in his heraldic coat can so stagger courage as the white-shrouded bear or shark.\*

\* With reference to the Polar bear, it may possibly be urged by him who would fain go still deeper into this matter, that it is not the whiteness, separately regarded, which heightens the intolerable hideousness of that brute; for, analysed, that heightened hideousness, it might be said, only arises from the circumstance, that the irresponsible ferociousness of the creature stands invested in the dress of celestial innocence and love; and hence, by bringing together two such opposite emotions in our minds, the Polar bear frightens us with so unusual a

Bethink thee of the albatross, whence come those clouds of spiritual wonderment and pale dread, in which that white phantom sails in all imaginations? Not Coleridge first threw that spell; but God's great, unflattering laureate, Nature.\*

Most famous in our Western annals and Indian traditions is that of the White Steed of the Prairies; a magnificent

contrast. But even assuming all this to be true, yet, were it not for the whiteness, you would not have that intensified terror.

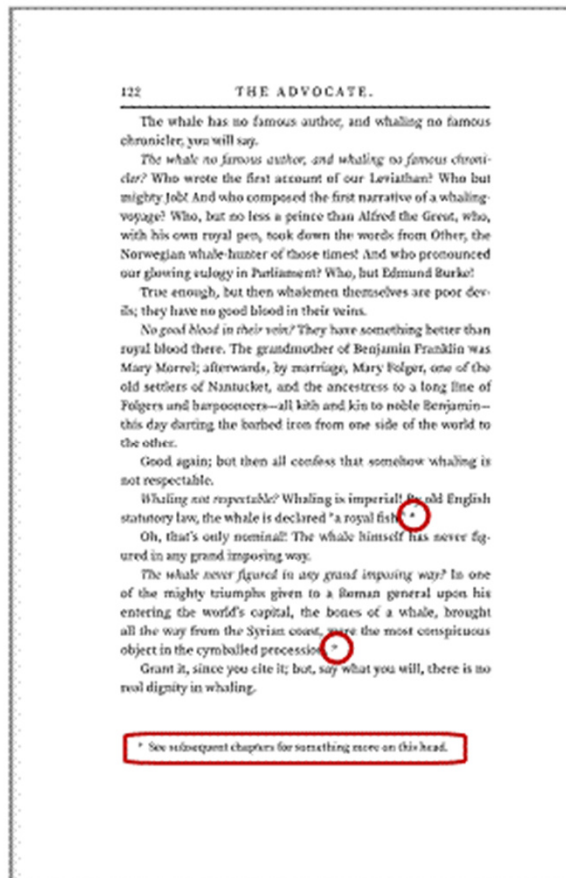
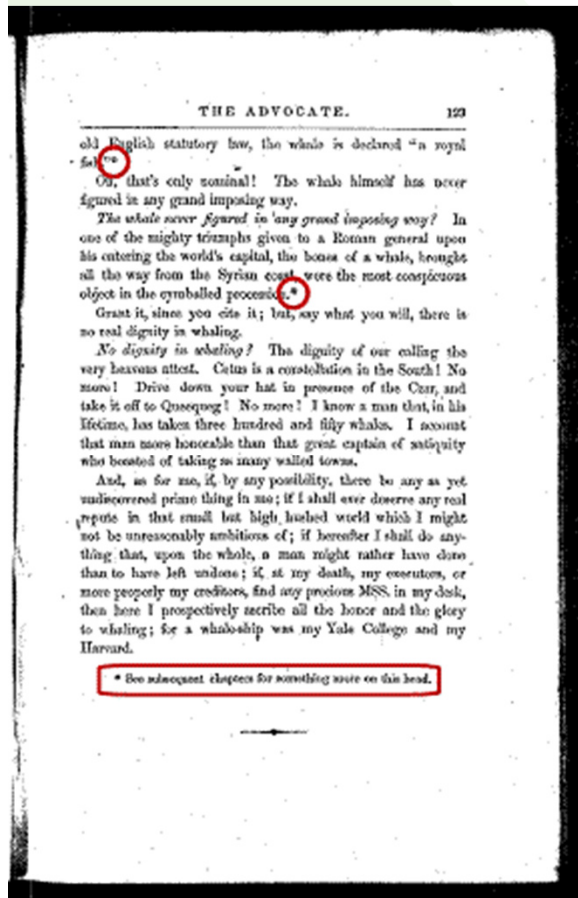
As for the white shark, the white gliding ghostliness of repose in that creature, when beheld in his ordinary mood, strangely tallies with the same quality in the Polar quadruped. This peculiarity is most vividly hit by the French in the name they bestow upon that fish. The Romish mass for the dead begins with "Requiem eternam" eternal rest; whence Requiem denominated the mass itself, and any other funeral tonic. Now, an allusion to the white, silent stillness of death in this shark, and the mild deadliness of his habits, the French call him *Requiem*.

\* I remember the first albatross I ever saw. It was during a prolonged gale, in waters hard upon the Antarctic seas. From my forenoon watch below, I ascended to the overclouded deck; and there, dashed upon the main hatch, I saw a regal, feathery thing of unspotted whiteness, and with a hooded, Roman bell sublime. At intervals, it arched forth its vast archangel wings, as if to embrace some holy ark. Windrows flutterings and throbbings shook it. Though bodily unharmed, it uttered cries, as some king's ghost in supernatural distress. Through its unexpressed, strange eyes, methought I peeped to secrets which took hold of God. As Abraham before the angels, I bowed myself; the white thing was so white, its wings so wide, and in those far ever chilled waters, I had lost the miserable weeping memories of traditions and of towns. Long I gazed at that prodigy of plumage. I cannot tell, can only hint, the things that darted through me then. But at last I awoke; and turning, asked a sailor what bird was this: A goney, he replied. Goney! I never had heard that name before; is it conceivable that this glorious thing is utterly unknown to men ashore? Never! But some time after, I learned that goney was some seaman's name for albatross. So that by no possibility could Coleridge's wild Rhymer have had sight to do with those mystical impressions which were mine, when I saw that bird upon our deck. For neither had I then read the Rhymer, nor knew the bird to be an albatross. Yet, in saying this, I do but indirectly burnish a little brighter the noble merit of the poem and the poet.

I assert, then, that in the wondrous bodily whiteness of the bird, chiefly lurks the secret of the spell; a truth the more evinced in this, that by a selection of terms there are birds called grey albatrosses; and these I have frequently seen, but never with such emotions as when I beheld the Antarctic fowl.

But how had the mystic thing been caught? Whipped it not, and I will tell, with a treacherous hook and line, as the fowl floated on the sea. At last the Captain made a postman of it, tying a lettered, leathern tally round its neck, with the ship's time and place; and then letting it escape. But I doubt not, that leathern tally, meant for man, was taken off in Heaven, when the white fowl flew to join the wing-folding, the invoking, and adoring cherubim!

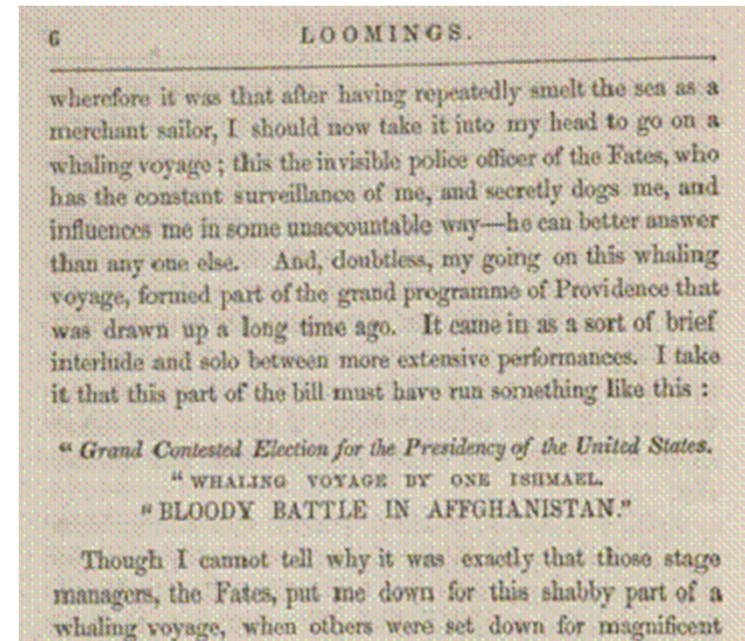
# 同じページに脚注が重複



- XML は脚注を繰り返す
- XSL 1.1 では統合されない
- `axf:suppress-duplicate-footnote`

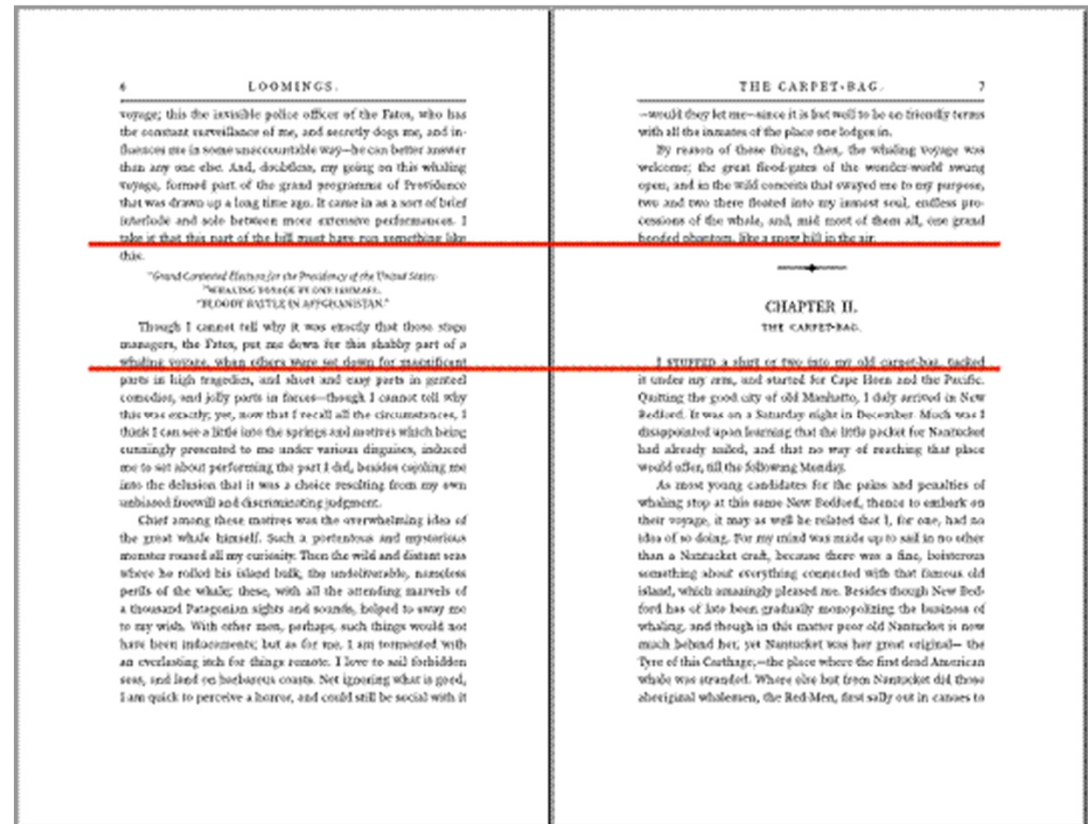
# 行グリッド

- 「裏写り」により可読性が低下
- 行の前後を揃える
- 不揃いな行やタイトルなどの後に再開する



# axf:baseline-grid

- 共通グリッドを設定
- ブロックに独自のグリッドを設定することも、全く設定しないこともできる



# 共通または特定のグリッド

## fo:page-sequence

```
<xsl:template match="body">
```

```
  <fo:page-sequence  
    master-reference="PageMaster"  
    initial-page-number="1"  
    axf:baseline-grid="root">
```

```
    <xsl:call-template
```

```
      name="static-content" />
```

```
    <fo:flow
```

```
      flow-name="xsl-region-body"
```

```
      hyphenate="true"
```

```
      text-align="justify">
```

```
        <xsl:apply-templates />
```

```
    </fo:flow>
```

```
  </fo:page-sequence>
```

```
</xsl:template>
```

## fo:block

```
<xsl:template match="body//q">
```

```
  <fo:block
```

```
    text-align="center"
```

```
    text-indent="0"
```

```
    space-before="0.251h"
```

```
    font-size="7pt"
```

```
    line-height="9pt"
```

```
    axf:baseline-block-snap=
```

```
      "before margin-box"
```

```
    axf:baseline-grid="new">
```

```
    <xsl:apply-templates />
```

```
  </fo:block>
```

```
</xsl:template>
```

# 引用の扱い

<p>"Do you is all sharks, and by natur wery voracious, yet I zay to you, fellow-critters, dat dat voraciousness—'top dat dam slappin' ob de tail! How you tink to hear, 'spose you keep up such a dam slappin' and bitin' dare?"</p>

"Do you is all sharks, and by natur wery voracious, yet I zay to you, fellow-critters, dat dat voraciousness—'top dat dam slappin' ob de tail! How you tink to hear, 'spose you keep up such a dam slappin' and bitin' dare?"

"Do you is all sharks, and by natur wery voracious, yet I zay to you, fellow-critters, dat dat voraciousness—'top dat dam slappin' ob de tail! How you tink to hear, 'spose you keep up such a dam slappin' and bitin' dare?"

- 一般的でない表現は一般的な表現に合わせる

# 連続するemダッシュ

- 複数の emダッシュ
- フォントにスペースが含まれている
- 単一の文字に置き換え

```
<xsl:variable
  name="text"
  select="replace($text, '—&—&—', '⸻')"
  as="xs:string" />
<xsl:variable
  name="text"
  select="replace($text, '—&—', '⸺')"
  as="xs:string" />
```

—— “There Leviathan,  
Hugest of living creatures, in the deep

— “There Leviathan,  
Hugest of living creatures, in the deep

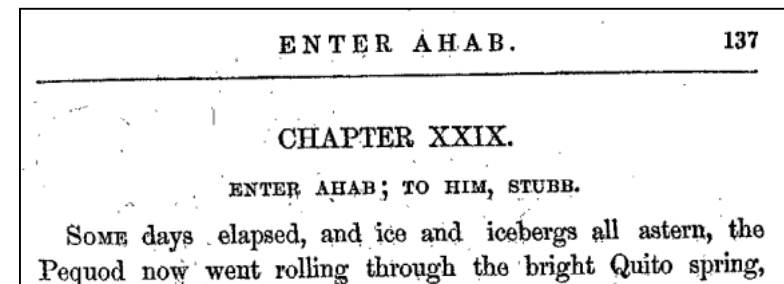
— “There Leviathan,  
Hugest of living creatures, in the deep

# ヘッダ

```
<div type="chapter">  
  <head>CHAPTER XXIX.</head>  
  <head type="sub">ENTER AHAB; TO HIM, STUBB.</head>  
  <fw type="head" place="top-centre">ENTER AHAB.</fw>
```

- `<fw>` が存在する場合は、その要素を選択する

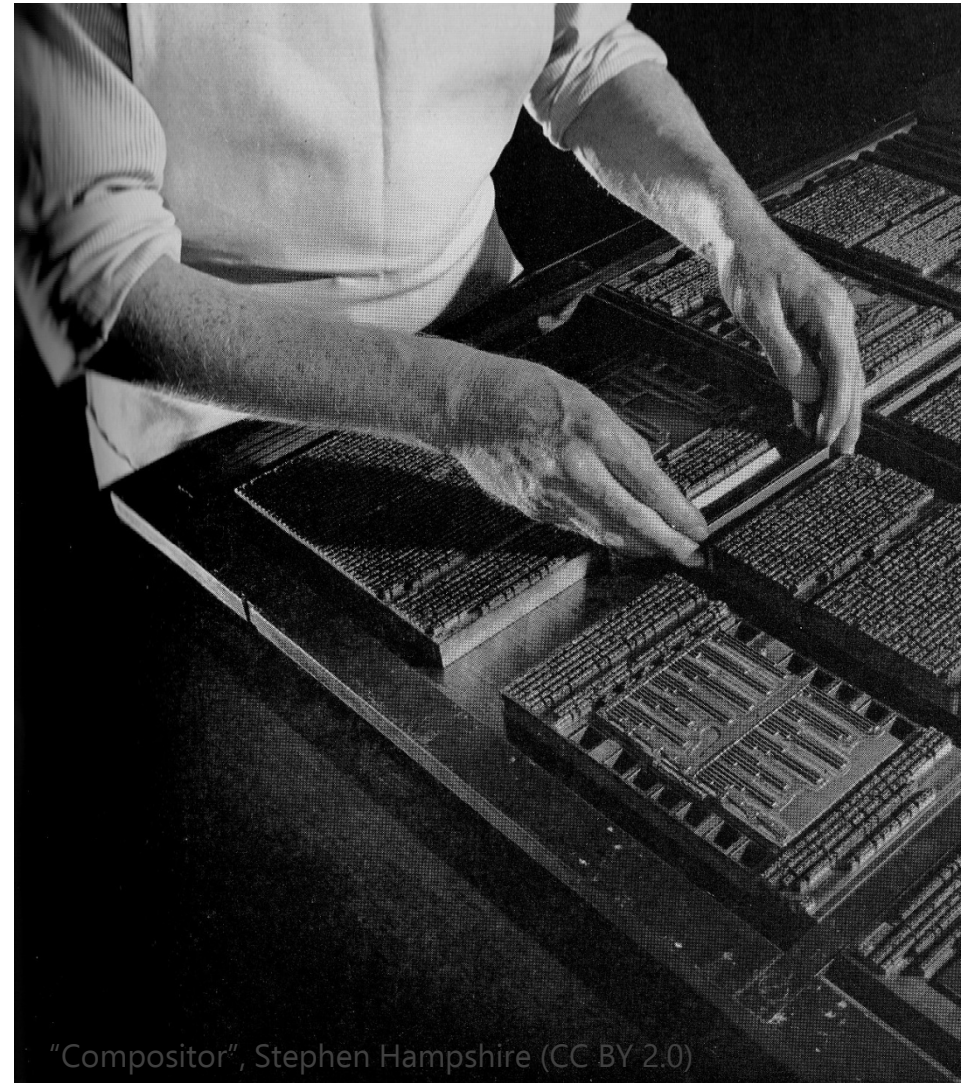
```
<fo:marker marker-class-name="Chapter-Title">  
  <xsl:apply-templates  
    select="(fw[@type = 'head'], head)[1]/node()"  
    mode="marker" />  
</fo:marker>
```





# Forme work

- <fw> は “forme work” のこと
- forme: 印刷用に取得した活字本体



# ヘッダ

- ヘッダは中央に配置
- 例外もある
- 長いヘッダをページ番号から遠ざけるには

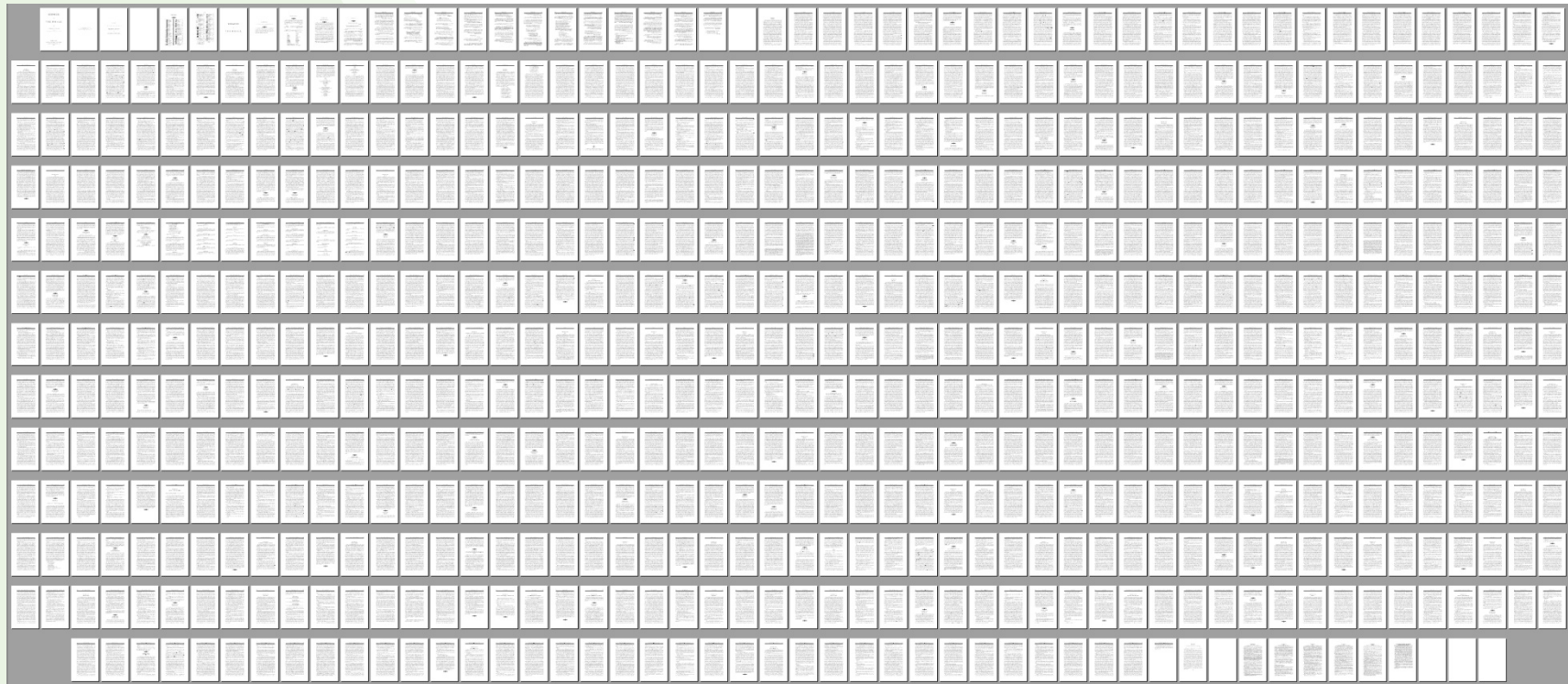
**axf:overflow-align="end"**

ENTER AHAB.	137
CHAPTER XXIX.	
ENTER AHAB; TO HIM, STUBB.	
Some days elapsed, and ice and icebergs all astern, the Pequod now went rolling through the bright Quito spring,	

THE HONOR AND GLORY OF WHALING. 405
up by a whale; still, whether that strictly makes a whaleman

THE HONOR AND GLORY OF WHALING. 405
brawny doer of rejoicing good deeds, was swallowed down

# Moby-Dick; or, The Whale(白鯨)



# まとめ

- まだ開発中である
- Moby-Dick stylesheet  
<https://github.com/AntennaHouse/moby-dick>
- Moby-Dick TEI XML  
<https://github.com/AntennaHouse/moby-dick-tei>
- 組版結果の自動分析機能は優れたもの  
<https://www.antenna.co.jp/AHF/help/en/ahf-analyzer.html>



*A Data Usability Company*  
**ANTENNA HOUSE**

ご質問はありますか？

**Tony Graham**

XML Division

Antenna House, Inc.

tgraham@antenna.co.jp

tony@antennahouse.com

@tgraham\_antenna

# 参考文献

- Moby-Dick stylesheet  
<https://github.com/AntennaHouse/moby-dick>
- Moby-Dick TEI XML  
<https://github.com/AntennaHouse/moby-dick-tei>
- Antenna House Formatter  
<https://www.antenna.co.jp/AHF/>
- Antenna House Formatter Automated Analysis  
<https://www.antenna.co.jp/AHF/help/ja/ahf-analyzer.html>

# Resources

- IU Digital Library Program. Moby-Dick, or, The Whale. Melville, Herman, (1819–1891).  
<http://webapp1.dlib.indiana.edu/TEIgeneral/view?docId=wright/VAC7237&brand=wright>
- IU Digital Library Program. Moby Dick, or, The Whale.  
<http://dogwood.dlib.indiana.edu:8080/xubmit/rest/repository/wright/VAC7237.xml>
- IU Digital Library Program. Wright American Fiction. Indiana University.  
<http://webapp1.dlib.indiana.edu/TEIgeneral/welcome.do?brand=wright>

# Resources

- “Compositor”, Stephen Hampshire  
<https://flickr.com/photos/stephenhampshire/3337704597>
- CC BY 2.0  
<https://creativecommons.org/licenses/by/2.0/>