

# Part1

## コンテンツとは何か？

Chapter 1. データ、情報、コンテンツを定義する

Chapter 2. コンテンツにはフォーマットがある

Chapter 3. コンテンツには構造がある

Chapter 4. 機能もコンテンツだ！

Chapter 5. コンテンツとは一体何なのか？

## データ、情報、コンテンツを定義する

### POINT

- ▶ データとは何か
- ▶ データとコンテンツはどう違うのか
- ▶ 情報をコンテンツに変える目的とは
- ▶ データを加えて情報をコンテンツに変えるには

本書の導入にあたって、なぜコンテンツの概念などという長々しく込み入った議論から始めるのかと、疑問に思っている人も多いかもしれない。それをまず説明しよう。私が知っているほとんどの組織は、極めて単純な理由から、コンテンツマネジメントのことを考えるようになっていく。ウェブや他の媒体で使いたい情報が膨大な量に上っており、それを助けてくれるツールがほしい、という理由だ。当然のことながら、彼らは答えを探している。あるいは少なくとも、コンテンツマネジメントを実践して、組織の本業に取りかかれる状態にしてくれる「ベストプラクティス」を探している。けれども残念ながら、ほとんどの組織は、答えを見つける代わりに、どこまでも続く疑問の糸を見つけてしまう。この長い糸をたどっていくと（通常は、いらだたしいミーティングを何度も繰り返すのがこの過程だ）、やがて組織がほどこかれていく。この糸は、組織のあらゆる部門に巻きついていて、編集や公開のプロセスすべてに絡まっている。この恐ろしいプロセスが解き明かされるにつれ、プロジェクトの範囲とコストは倍々に膨れ上がり、逆にチームの士気と情熱は急降下する。

私の考えでは、こうした組織が抱えている根本的な問題は、データとコンテンツの違いを誤解していることにある。コンテンツに対して、データのようなシンプルさ、分かりやすさを期待しているのだ。

本書のPart1では、コンテンツという概念を詳しく考察していきたい。コンテンツは、なぜコンピュータがこれまで処理してきたデータと異なるのか。コンテンツは、なぜシンプルでもなければ分かりやすくもないのか。これらの疑問について、確かな答えを示していきたい。コンテンツの概念をしっかりとらえることができれば、多くの組織がはまる泥沼を回避できる（あるいは抜け出せる）と、私は考えている。自社の組織を這い回っているコンテンツの糸をほどこうとして、逆に糸に絡まり、方向が分からなくなってしまうのではなく、この糸を最初から順にたどっていくのだ。もちろん、糸が途中で絡まっていることもあるだろう。しかし、たくさんの結び目をほどこく準備は整えられる。さらに重要な点として、この混乱に対処する間、現実的な目標を見失わず、方向感覚を保てるようになるだろう。

コンテンツの本質を理解することは、もちろん、コンテンツマネジメントのスタートライ

ンにすぎない。これから立ち向かおうとしている相手が誰なのかを、まず頭のなかで理解したうえで、そこから実践可能なプロジェクトを策定し、自分の置かれた特有の状況を分析し、そして最後に、組織が収集するコンテンツを消費する必要がある人に届けるシステムを構築しなければならない。これらの内容は、本書全体を通じた後半で取り上げていく。

## 1. データとは何か

コンピュータは、「データ」を処理するために作られている。データは、数値、言葉、画像、サウンドといったコンピュータ情報の細かい断片で成り立っていて、その情報からは、人間に意味のあるものは、ほとんど削ぎ落とされている。今日、人はコンテンツを処理するのにコンピュータを使っている。データと同じように、「コンテンツ」も情報だが、人間にとっての意味やコンテキストがあるという点は大きな違いだ。

このChapterでは、コンテンツマネジメントが抱える基本的な問題の1つを説明していきたい。それは、コンピュータは「データ」を処理する目的でデザインされている、という問題だ。データには、コンテキストもなければ、それ単体での意味もない。けれどもユーザは、コンテキストも意味もいっばいの「コンテンツ」を、コンピュータに処理してほしいと思っている。データテクノロジーを使って、まるでデータとは異なるコンテンツを、どうすれば管理し、提供できるのか。これはやさしい課題ではない。情報を強引にデータのように仕立ててしまえば、機械的で味気なく、消費する人にとって理解しにくいものになってしまう。逆に、情報をあまりにもリッチで変化に富み、コンテキスト満載のものにしてしまえば、コンピュータに自動管理させることができなくなる。

その妥協点は、本書でも紹介していく通り、情報をデータを格納する入れ物（データコンテナ。「メタデータ」として知られている）に入れてやることだ。コンピュータはデータを処理し、その結果、意味のある面白い情報が届けられるようになる。

もともとコンピュータは、人間がやるにはあまりにも複雑で時間のかかる計算をするために発明された。根底にあるモデルは、次のようなものだった（今でもその多くは変わっていない）。数値や論理要素（真か偽かを示す要素）を使ったシンプルで機械的なオペレーションであるアルゴリズムにまで問題を落とし込むことができれば、コンピュータによる解決が可能となるというモデルだ。デイヴィッド・パーリンスキ氏は、そのすばらしい著書『史上最大の発明アルゴリズム』のなかで、アルゴリズムについて次のように説明している。

アルゴリズムは  
有限の手続きであり、  
定まった一組の記号で書かれ、  
厳密な指令に支配され、  
1、2、3・・・ という個々のステップを踏んで進み、  
これを実行するのに、洞察力も才気も直観も知性も明敏さも必要なく、  
遅かれ早かれ、終わりにいたる——そのようなものである。

(邦訳『史上最大の発明アルゴリズム』より引用、早川書房刊、2001)

黎明期のコンピュータ専門家には、プログラマかデータ入力担当者のどちらかしかいなかった。プログラマは、問題を機械的なオペレーションに落とし込む。それを支えるのは、次のようなシンプルな格言だ。

データを入力すれば、コンピュータが処理して、もっと使いやすい形式でアウトプットしてくれる。

データ入力担当者は、データ入力だけが仕事だ。縦横に規律正しく机を並べて、これまた縦横に規律正しく並んだ数字や、苗字、名前、住所などの短いフレーズをタイプしていく。やがて時代が進み、コンピュータ科学の研究者がデータベースを発明して、膨大な量になったこれらの断片を整理・保管できるようにした。

お分かりの通り、世の中にある様々な問題のなかには、こうやって解決したほうがいいものがある。こうして、コンピュータ技術が発展するにつれて、コンピュータは科学研究の分野だけでなく、製造や財務会計でも使われるようになった。大量の数値を管理することが、これらの分野でも大きな部分を占めていたからだ。もちろん、今ではコンピュータがすべてのものに使われている。機器、ツール、おもちゃ、半導体プロセス、エンタテインメント、そして言うまでもなく職場で。しかし、このすべてのものについてコンピュータが進入したのは、軽減することのできる部分だ。軽減できる部分とは、数値と論理要素によって動く極めて限定的な規則の有限の組み合わせが、有用な結果を生み出すことのできる部分のことである。

---

## 2. コンテンツはデータではない

データ処理マシンとしてのコンピュータという発想は、非常に根が深い。今日に至るまで、データ処理は、コンピュータがこなす仕事の主な部分を占めている。一方で、誰もが知っている通り、ほとんどのユーザは、断片の山を細かく砕く以上のことをコンピュータにやってほしいと思っている。コンピュータが、完全な状態の巨大な情報のかたまり(断片ではなく)の山をふるいにかけて、その瞬間に一番ほしいと思うものを提供してくれることを、今では多くの人が望んでいるのだ。さらに、本やラジオ、テレビ、映画のような、よりなじみのある情報ソースに期待するクオリティでの情報提供もコンピュータに望んでいる。

コンピュータは、製造や財務会計の域をも飛び出し、職場の机の上で、タイプライタや紙ベースのスプレッドシート（集計表）に取って代わった。この時、パーソナルコンピュータ業界では、3つの関連した発展があった。それらのブレークスルーが一緒になって、コンピュータができること、コンピュータがすべきことに対する私たちの期待に、大きな変化の起きる下地を整えた。

- デジタルメディアの作成（画像、サウンド、ビデオ）が可能になった。
- デジタルメディアの出力機器（カラーモニター、サウンドカード、ビデオアクセラレータ）が出回るようになった。
- 一般消費者向けの大容量記憶装置（言い換えればCD-ROM）が出回り、値段も安くなった。

これらの発展によって、マルチメディア産業が彗星のごとく登場し、既存の業界でもマルチメディアが使われるようになった。ここへきて初めて、誰もが手軽に、単なるデータではない本当のコンテンツを作って提供できるようになったのだ。マルチメディアCD-ROMは瞬く間に普及して、百科事典やカタログからフルモーションゲームまで、あらゆるものを提供するようになった。コンピュータは、本、テレビ、ラジオといった伝統的な情報チャンネルの代わりとして浮上しはじめた。そう、これらの伝統的なチャンネルが提供していたのは、「コンテンツ」であって、データではなかったのだ。

マルチメディア産業が始めたことを、今、ウェブが完結しようとしている。今日、コンテンツをオンラインで手に入れることは、可能なだけでなく、多くの場合そのほうが都合が良いという状況になっている。私自身、今では音楽やトークをラジオやステレオで聴くよりも、ラップトップコンピュータやPDAで聴くことのほうが多い。それでもなお、ウェブにアクセスすると、たいていはフラストレーションが溜まる。ほしいコンテンツをすばやく見つけて、少なくとも伝統的なチャンネルと同じぐらいのレベルで提供してほしいのに、残念ながら必ずしもそうではないからだ。



## ノート

デジタルコンテンツの仕事に多少でもかかわったことがある人なら、ここに大きな逆説があることが分かるだろう。ユーザは、アクセスが簡単で、プレゼンテーションが魅力的であることを期待している。しかし、もしそれが達成されているとしたら、簡単に魅力的なように見せるために、背後で大変な努力を重ねた人のおかげだ。コンテンツを自然に見せるという作業は、現実には極めて不自然で難しい試みである。

こうして、ユーザのニーズや期待は変わったが、コンピュータの根本は変わらなかった。10年前、ほとんどの人はデータを入力し、処理し、出力するのにコンピュータを使った。今ではほとんどの人がコンテンツを見つけ、消費するためにコンピュータを使っている。にもかかわらず、すべてのコンピュータ技術の根底に依然として存在するのは、どんな問題でも個々のステップを踏んで動くシンプルな命令群と構造化された断片に落とし込む、という概念だ。

データとコンテンツは異なる。これは紛れもない事実だ。けれども、違いがあるからといって相互に影響しないということではない。実際、片方から他方への移行は、毎日数え切れな

いほど起こっている。もっと言うなら、コンピュータシステムという観点から見れば、コンテンツなどそもそも存在せず、データだけが存在するのだ。今日、ユーザは、コンテンツをコンテンツとして扱うツールをほとんど持ち合わせていない。だから、コンテンツをデータとして扱うことで、コンピュータが保存し、取り出し、表示できるようにしなければならない。

例えば、音楽サイトにアクセスした時の典型的なウェブインタラクションを考えてみよう。好きな音楽CDをフィーチャーしているページをブラウザして、ショッピングカートに入れ、代金を支払う。ここで体験するのは、音楽についての情報のほか、それを購入するのに使うボタンやその他のコントロールなどで構成された複数のウェブページだ。全体として、このユーザエクスペリエンスは、コンテンツリッチなインタラクションのように感じる。けれども、背後で起こっていることと言えば、データ志向のコンピュータプログラムがデータベースとデータを交換しているだけ、ということになる。

背後でやり取りされるデータのなかには、コンテンツに非常に近いものもある。例えば、データベースに保存されているものには、アーティストの写真を伴った特集記事などがある。あるフィールドにアーティスト名が、別のフィールドに記事のテキストが、さらに別のフィールドに写真が、というかたちで保存されている。データのなかには非常にデータらしいものもある。これらは、ユーザとレコード会社の間の決済手続きを成立させる、数字やその他の情報の断片で構成されている。データベースは、ユーザのクレジットカード番号、注文番号、注文個数、注文価格などを保存して、計算や他のアルゴリズム処理に使用する。一方、データとコンテンツの間に位置するようなデータもある。音楽カタログには、楽曲名、再生時間、価格、在庫の有無などが含まれている。こういった情報の断片は、ウェブサイトのデータベースが保存する時には極めてデータらしく見えるが、サイトに表示される時にはコンテンツらしく見える(図1参照)。

データベースの中では、すべてがデータにすぎない。ページ上では、決済データはなおも多

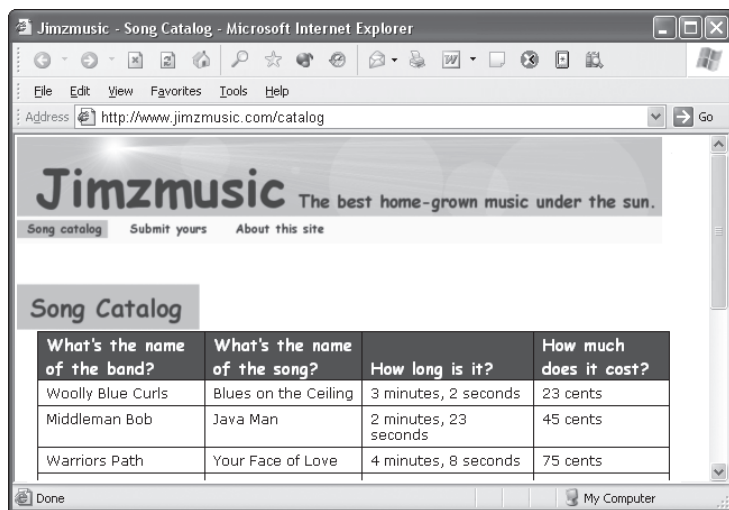


図1：サイトがコンテンツを格納する際には、他のデータとまるで変わりがなく見える。けれども、オーディエンスを獲得しようと思ったら、そのコンテンツを表示する際には、データのように見せてはいけない。

## 情報という色眼鏡で世界を眺めてみると

本当に、すべては情報に尽きる。仕事、学校、そして遊びまで、あらゆるタイプの人間の営みで、情報は中心的な役割を果たしている。情報科学の研究者は、世の中を次のように見ている。情報処理や情報システムが、様々な種類の情報を処理して、何らかの種類のニーズを満たしている――。

情報という観点は、様々なレベルで用いられる。生命を築く基本的断片であるDNAはマイクロレベルの情報システムであり、ニューヨーク証券取引所はマクロレベルの情報システムである、という具合だ。

情報を認識するうえでは、「情報スペクトル」というものが土台となる。データから始まり、情報、知識、行動へとわたるスペクトルだ。データとは、未加工の記号、数値、文字、例えば「30.41 - 45.99 APA 6-30」といった、モノの基本的な表現形態だ。データは、ラベルを与えられた時点で意味を持つようになる。例えば、「52週間、記号、日付」などだ。知識は、たくさんの情報を集めて合体させ、合成し、結論を導くことによって生じる。例えば、「アパッチ株式会社の株価は昨年50%上昇し、APAは5月に配当金を出した。これを受けて6月1日、ファースト・オルバニーのアナリストは同社の格付けを『買い』に引き上げた」などだ。そして行動は、その情報に基づいて何かすること、つまり、買う、売る（現状維持でもいい）などの行動を取ることを意味する。

もちろん、知識には至っていないと判断して、データや情報の収集に戻って、より良い行動が取れるようにすることもできるかもしれない。

情報スペクトルは、業務上のことであれプライベートなことであれ、人間が置かれるあらゆる状況で発生する過程を表している。今度、仕事上の決定を下したり、クラスを受講したり、はたまた新聞を読んだり、野球の試合を観たりすることがあったら、情報スペクトルのことを思い出してみよう。

すべては情報に尽きる。

ワシントン大学情報学部 ([www.ischool.washington.edu](http://www.ischool.washington.edu))

学部長

マイケル・B・アイゼンバーク

少しはデータのように見えるが、特集記事のデータはまるでデータのように見え、またカテゴリーデータは、どこまでデータらしく見せるかを加減できる。とはいえ最終的にページのデザインがうまく作られていれば、ビジターはすべてをコンテンツと認識するだろう。

つまり、ユーザの見地に立てば情報はすべてコンテンツであり、コンピュータプログラマの見地に立てばすべてデータということだ。すべてのテクノロジーがデータ志向である今の時代では、データテクノロジーを使ってコンテンツの保存と表示を行うことが、コンテンツマネジメントの秘訣ということになる。

### 3. コンテンツは使うための情報だ

コンピュータの世界に携わっている人は、多かれ少なかれ、「データ」の定義について一致した考えを持っている。データとは情報の小さな断片で、人はこれを収集し、結合させてデータレコードを形成し、データベースに保存する。一方、「情報」という言葉には、たくさんの意

味が含まれると同時に、まったく意味がないこともある。情報を含めどんなものであれ、データと呼ぶことができるからだ。

「情報」という言葉には多くの意味がある。本書の目的上、私はかなりありふれた定義を用いることにしよう。記録されたコミュニケーションの一般的な形態すべてを「情報」と呼ぶことにする。具体的には、次のようなものが含まれる。

- 記事、本、ニュースなどのテキスト
- 音楽、会話、朗読などのサウンド
- 写真、イラストなどの画像
- ビデオ、アニメーションなどの動画
- スプレッドシート、スライドショー、そのほかユーザが探して使いたいと思う独自形式の書類などのコンピュータファイル

情報は、誰かの目に触れる前に、必ず別の人がたくさんの労を費やして作った経緯を持つ。情報を作った人は、伝えたいコンセプトのイメージを心のなかで描き、創造力や知性を駆使して、コンセプトにふさわしい言葉やサウンド、画像に仕立て上げたのだ（こうして、コンセプトを明確化した）。それからこの人は、その情報を何らかの提供可能な形式で記録した。情報の作者は、それを誰かが目にするよりも前に、情報に対してたくさんの個性とコンテキストを注ぎ込んでいる。このため情報は、データとは違って、最初から分かりやすい小分けの容器に入っていて、どれも同じ構造で、同じように動作する、というものではない。情報は、ちょうど人の会話のように、途切れることなく流れつづけ、標準的な開始時点や終了時点、そして属性などもないことが多い。この連続性を分断する人は、自らの意思で危険を冒すことになる。情報を分断ということは、情報を変更することを意味する。そして、作者がその情報で表現したいと思ったオリジナルの知性や創造性を、変えてしまったり失ってしまったりするリスクを負う。また、情報を分断すれば、作者がオーディエンスや情報の目的に対して仮定していた意図も、分からなくなったり無視してしまったりするリスクを負う。

かつて存在したContentWatchという組織（コンテンツマネジメントの理論と実践を研究するグループだった）は、コンテンツの定義について、次のような有意義な見方を打ち出していた。

「コンテンツ」とは何か。生の情報は、1つ以上の目的のための有効な形態を与えられた時、コンテンツになる。コンテンツの価値は、その主な有効形態と、用途、アクセシビリティ、使用状況、有用性、ブランド力、独自性のコンビネーションによって決められるようになりつつある。

世の中を何気なく通過していく情報は、コンテンツではない。誰かがそれをつかみ、何らかの目的に役立てようとして初めて、コンテンツになる。情報をつかんで役立つようにする方法には、情報の周りにデータの層を加えるやり方がある。この際のデータは、通常、その情報を説明し、コンテキストを与えるものだ。これにより情報は、コンピュータにとっても人間にとっても使えるものになる。簡単な例を挙げるとすれば、このChapterのタイトルは、一片



の「データ」だ。そして、すべてのChapterのタイトルと見出しを集めれば、目次と呼ばれる「情報」になる。この情報を本の冒頭に入れれば、読者はそれを使って本をナビゲートし、何か価値のあるものを見つけられるようになる。これが「コンテンツ」だ。

情報にデータを付け加えることは小さなステップのように見えるかもしれない。情報からその使用方法へと進むだけのことだ。けれども、これは決して小さなステップではない。情報の本質からその実践的な使用方法にフォーカスを移すことによって、データに対する考え方を情報に適用するという、まったく新しい可能性を持つ世界が開けるのだ。データについて考えるときのポイントは、情報をデータのように扱うことはできなくても、情報の「使用方法」をデータのように扱うことはできる、ということだ。情報を使い始めるにあたって、情報を単純化する仮説(メタデータ)で包み込めば、コンピュータを使って、その使用法を管理できるようになる。情報を作り、それを単純化する仮説を考え出すうえで人間の存在が欠かせないが、そこから先はコンピュータで処理でき、たいていは意味の通ったやり方で、情報を

## 情報、変化、そして情報プロダクト

私の情報の定義は、本書の著者であるボブ・ボイコ氏の定義を補完するものだ。

情報とは、人間が知識をもって他の人とコミュニケーションしたい時に変化させ、その結果となるものを言う。見えるように、あるいは聞こえるように変化した知識、または、書かれた言葉、印刷された言葉、話された言葉で表される知識が、情報である。

ユーザの側からすれば、情報とは、何か行動を起こすために知識を増やしたり深めたりする必要がある時に、意識的に外界に求めるものである。そして情報を見つけると、私たちはそれを知識に変化させる(※)。

ボブが言うには「どんなものであれ、データと呼ぶことができる」。この「どんなものであれ」を情報に変えるのは、自分の知識にそれを加えて使いたいと思う人にとっての重要性だ。

知識を情報に変化させるのは、「記録されたコミュニケーションの一般的な形態すべて」が世に出る前に、誰かが費やした「たくさんの労」だ。この結果としてできあがったものを、私は情報プロダクトと呼んでいる(自分でもあまり好きな言葉ではないが、今の時点で使える唯一の言葉だ)。そしてこれを、私は次のように定義している。

情報プロダクトとは、紙の上の印刷であれ電子的なものであれ、目に見える(または耳で聞こえる)プロダクトであり、それを通じて情報が使用目的のために提供されるプロダクトである。情報プロダクトは、知識が情報に変化した結果を具現化し、コンテンツとその入れ物が分離できない状態に一体化したものを指す。

人間の知性と、情報プロダクトに具現化された情報の出会いが、この結果をもたらす。

オーナ・インフォメーション&エディトリアル・コンサルタンシー  
リズ・オーナ (lo@orna.co.uk)

※ E・オーナ著『Information Strategy in Practice (実践的情報戦略)』  
(2004年、英国オルダーショット、Gower刊) p.7より

分配してくれるようになる。先の目次の例を使うならば、私が、タイトルや見出しについての仮説を作らなければならない。つまり、タイトルと見出しは本の情報を整理するためのものである、タイトルと見出しはセクションの内容を簡単に示すためのものである、といった仮説だ。けれども、タイトルや見出しを集めて読みやすいフォーマットに整え、ページ番号が正しいかどうかを確かめるといった作業は、ほとんどコンピュータに頼ることができる、ということだ。

情報をデータで包むことによって、ある1人の人間による小さなアクションが、コンピュータに対してたくさんの動作を実行させるためのトリガーとなる。例えば、情報管理を簡略化するために、情報を(1)新規、(2)発行待ち、(3)削除待ちの3種類に分けることにしたとする。コンピュータは、どの情報にどのステータスが当てはまるのかを決めることはできない。けれども、情報を「ステータス」というメタデータで包み、人間がステータスのメタデータを設定することによって、そのステータスに基づいたたくさんの仕事をコンピュータに実行させることができるようになる。コンピュータは、情報自体について何も知っている必要はない。その情報に人間がどのステータスを付けたかだけを知っていればいいのだ。そして、動作は次のようになる。

- 人間がコンテンツの断片に「新規」のステータスを与えた場合は、コンピュータは指定されたレビュー担当者に標準のEメールメッセージを送信する。
- 人間がコンテンツの断片に「発行待ち」のステータスを与えた場合は、コンピュータはそれを指定されたウェブページにアウトプットする。
- 人間がコンテンツの断片に「削除待ち」のステータスを与えた場合は、コンピュータはそれをウェブページから削除して、システムからも削除する。

このように、人間が施すちょっとした作業の結果として、コンピュータがたくさんの作業を行えるようになる。

オンライン版『メリアム・ウェブスター・ディクショナリ』(メリアム・ウェブスター・オンライン [www.m-w.com](http://www.m-w.com))では、コンテンツの定義の一部として、次のような説明を掲載している。

1 a: 中に入っているもの。通常は複数形で用いられる。[the jar's contents (瓶の中身)] [the drawer's contents (引き出しの中身)] b: 書物で取り扱われているトピックやものごと。[table of contents (目次)]……

この定義は、コンテンツを理解するにあたって有益な見方をもたらしてくれる。「何かの」中に入っているもの、という見方だ。つまり、情報をコンテンツに変えることによって、あるモノに対する関心をモノを入れる容器のほうに移すことができる。焦点が、情報そのものではなく、それを取り巻いているメタデータのほうに移動するのだ。情報を入れる容器とは、カテゴリやメタデータの組み合わせであって、そう、情報はその中に入っていることになる。付加されたデータが情報を囲い込んで閉じ込め、パッケージ化して、使用、再使用、目的の再設定、そして再配信などを可能にする。さらに入れ物の喩えを使って説明すると、コンテン

## コンテンツと入れ物

私は、メリアム・ウェブスター・ディクショナリにあるコンテンツの定義が非常に的確だと思っている。コンテンツの概念そのものが入れ物の存在を示唆することを、うまくとらえているからだ。しかし、入れ物とは何かという点について、私自身はボブよりも広義な見方をしている。コンテンツと入れ物が分離できない状態に一体化したものが情報プロダクトだと、私は考えているからだ。入れ物の内側の層は、思考を表現する言語（テキスト、ビジュアル、音楽など）で成り立っている。そのほか、コンテンツが構成され表現されている方法、ユーザやユーザの使用法、またコンテンツそのものにふさわしいプロダクトにするために、情報デザインという名の下にコンテンツに対して施されるすべてのことが、やはり入れ物の内側を作り上げている。そして、メタデータやカテゴリ、索引の用語などが一番外側の層を作り上げ、これが、製品を識別するためのラベルや、ユーザが必要に応じてコンテンツを見つげられるようにするためのラベルとなっている。

オーナ・インフォメーション&エディトリアル・コンサルタンシー  
リズ・オーナ (lo@orna.co.uk)

ツが市販の蜂蜜ならば、メタデータはその蜂蜜が入った瓶ということになり、瓶には、蜂蜜のブランドやグレード、容量、産地などを説明したラベルが付けられているだろう。

コンテンツが使うためにある情報だとすれば、最初に問うべき質問は、「何のために使うのか」ということになる。すべての情報を必死になって整理整頓する目的は、いったい何なのか。この問いかけは、極めてシンプルなものに聞こえるかもしれないが、実に多くのコンテンツマネジメントのプロジェクトが、この質問に答えないうままに進んでいることを考えてみてほしい。もっと正確に言うなら、すべてのプロジェクトが何らかの目的は持っている。けれども、その目的が、プロジェクトで扱われるコンテンツとほとんど関係がないのだ。「なぜこのウェブサイトを作るのか」という質問に対して次のような答えが返ってくるのが、あまりにも多い。

- 必要だから。
- 競合他社がやっているから。
- CEOが必要だと言うから。
- みんながほしいと言うから。

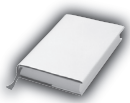
この種の答えは、プロジェクトのきっかけにはなるが、プロジェクトを定義するものではない。プロジェクトを定義するには、「まとめようとしているコンテンツの『目的』は何なのか」という質問に答えなければならない。これに対してよく練られた確かな答えを返せるのであれば、自然とコンテンツの整理方法も分かるようになり、目標は達成される。優れた目的を持つためのカギは、「具体的で測定可能な」目的とすることだ。例えば、イントラネットを作る際の目的としては、次のような模範回答がある。

- 生産部門が提供する新製品データを、現場の営業担当者が24時間いつでも取り出せるよ

うにする。

- 既知の競合製品について言及している、確実な情報ソースからの記事をすべて、サイトに掲載する。
- 給与明細に記載されたすべてのデータを、全従業員の完全な経歴情報とともに表示できるようにする。

ここに書かれたゴールは非常に具体的なため、その数によってはかなり大規模なイントラネットの必要性が浮き彫りにされるかもしれない。こうして、ゴールを整理していくうちに、その背後にあるコンテンツを整理していくことになる。個別のゴールがすべて、一貫性のある組織全体のゴールにフィットするようになって初めて、整理のプロセスは完了し、最終的に、「現場をフルサポートする」「従業員の質問にすべて明確に答える」といったシンプルなゴール宣言文に落とし込むことができる。



### リファレンス

Chapter 16「プロジェクト委任を確保する」の「4-5. プロジェクト委任」のセクションでは、組織全体の目標をプロジェクトのゴール宣言文に落とし込むプロジェクト委任という考え方について説明している。

## 4. コンテンツは、情報にデータをプラスしたもの

情報には、どこか人間的で勤に頼るような部分があるために、データのように簡単に処理することが不可能となっている。データは、見た通りそのままのものだ。けれども情報の場合、それが伝えることのほとんどは、実際に情報のなかにあるのではなく、情報を作成する人や消費する人の頭のなかにある。情報は、言外の意味やコンテキスト、解釈といった広大な世界に存在していて、そのほとんどは、根本的にデータ処理のモデルでは扱うことができない。もともと、データの問題は、情報からこれらの主観的な性質を取り除いて、コンピュータが厳格な論理的正確さで管理できるようにするために、作られているのだ。

では、コンピュータが情報を管理できるようにするのは可能なのか。私は可能だと思っている。ただし、人間の基準から言うと、下手ながらも、というレベルだ。新しいコンピュータ処理のモデルが出現するまでは、方法論やプロセスや手順を開発することによって、データ時代のツールを使って情報を「扱う」ことはできる。情報を単なるデータに縮小してしまうのではなく、意味のあるかたまり全体としてとらえ、コンピュータが読んで実行できるように説明的なデータで包んでやるのだ。

メタデータは、データに関するデータと簡単に説明されることが多い。しかし、本当はそれ以上のものだ。メタデータによって、情報のコンテキストや意味が十分に明らかになるからこそ、コンピュータで扱えるようになる。

情報に付加されたデータ（つまりメタデータ）で、情報を全体として維持する行為と、データテクニックで効率よく管理する行為の間にあるギャップが狭まる。情報に付加するデータ

は、コンテキスト、言外の意味、そして解釈を明確にする手段だ。さらに、メタデータは、情報を解釈する人に対する期待も明確にする。例えば、作成する情報のかたまりそれぞれに「オーディエンスのタイプ」というメタデータを付加すれば、情報のターゲットに関する暗黙の仮定を明確にできる。ターゲットには例えば、「専門ユーザ、犬の飼い主、下請業者」などがあるだろう。そして、画面の向こう側に誰がいるかに応じて、コンピュータは、情報を探し出す簡単なタスクをこなせるようになる。もちろん、すべてのタスクが簡単というわけではない。が、概念は常に同じだ。人間がデータを用いて情報の大きなかたまりにタグを付けると、コンピュータはデータにアクセスして、情報をどう扱うかを決定する。メタデータは、情報のコンテキストや意味を、コンピュータが理解して作業できるレベルにまで単純化したバージョンということだ。

ゆえに、コンテンツとは、「人間がタグ付けすることによって、コンピュータがその収集と管理、そして発行の手続きを整理し体系化できるようになった情報」と言うことができる。「コンテンツマネジメントシステム (CMS)」は、データの方法論を適用しながら、その過程で情報の価値や意味をつぶしてしまわずにコンテンツを処理できるのであれば、うまく機能していることになる。

コンピュータ（あるいは何か別の新しいテクノロジー）がコンテンツを直接的に扱えるようになるまでは、データテクノロジーを使ってコンテンツを収集し配信する方法を考えなければならない。コンテンツを扱うのにデータテクノロジーを使用するというところこそが、本書の中心的なテーマなのだ。

## 5. データからコンテンツへ、そして再びデータへ

本書はデータとコンテンツの間の差を縮めることを目指しているが、その観点からすれば、何がデータで何がコンテンツなのかは、さまざまな種類の情報の作り方、管理しかた、ページ上での表示方法などによって大きく変わってくる。情報の種類には、次のようなものがある。

- **取引情報** 通常、物品の売買プロセスを追跡するために使われる、データに似た情報の断片を指す。取引情報やその他の極めてデータライクな情報素材に対しては、一般に、CMSを使うことはない。たいていは、従来型のデータ処理システムの一部として管理されるからだ。けれども多くの場合、CMSのテンプレートシステムでは、発行物のページ上に表示されるこの種の情報を管理するようになっている。
- **記事情報** テキストが多く使われ、何らかの編集プロセスが介入して作られた情報を指す。記事情報やその他の極めてコンテンツ性が高い情報は、ほとんどの場合、CMSで作成して管理する。実際、この種の情報がなければ、CMSの存在意義はないと言ってもいいほどだ。また、CMSは一般に、この種の情報の表示も司る。
- **カタログ情報** ディレクトリや製品リストとして表示されるような情報を指す。この種の情報は、CMSで作成、管理することもあれば、別のデータ処理アプリケーションで処理することもある。当然のことだが、すでに完全な機能を備えたカタログシステムがあるのであれば、それをCMSと接続するのが最も簡単なやり方だ。一方、部品番号や値段だけでなく、リッチメディアや大量のテキストをカタログに盛り込む必要があるのであれば、包括的なCMSの一環としてカタログ機能を構築するほうが、理にかなっている。

コンテンツマネジメントの目的は、すべてのデータをコンテンツに変換することではない。コンテンツマネジメントの目的とは、編集に重きが置かれるリッチな情報の作成と管理を監督し、この種の情報が既存のデータシステムに統合されるプロセスを管理することだ。どんな場合であれ、CMSは、最終的な発行物の生成を担わなければならない。この発行物がデータとコンテンツの両方を含んだものになるのであれば、CMSは、正しいデータと正しいコンテンツが正しい場所に表示され、その発行物が全体として統一感ある印象をエンドユーザに与えるものとして、その役割を果たすことになる。

## 6.まとめ

データは、情報よりも扱いがずっと簡単だからという理由で発明された。データは、小さくて、シンプルで、すべての関係性が明らかに分かっている（もしくは無視されている）。データがあるからこそ、コンピュータプログラムを書くことが可能になる。一方、情報は、大きくて、複雑で、その意味にとって重要な関係性に覆い尽くされているが、コンピュータが読み解くことはできない。

データ、情報、コンテンツという言葉を目にし、議論していくにあたって、心に留めてほしい点を、以下に示す。

- コンテンツは（少なくとも本書のなかでは）、データの有用さと情報のリッチさを足して二で割ったものだ。コンテンツとは、シンプルなデータに包まれたリッチな情報だ。情報を取り巻くデータ（メタデータ）は、情報のコンテキストや意味の単純化バージョンだ。
- CMSでは、コンピュータがデータを介して非直接的に情報を管理する。コンテンツは、おそらく妥協してこう言うだろう。「コンピュータに情報を理解して管理してもらうことができないから、問題を単純化しなければならないんだ。情報の重要な側面について、私ができる最高の推測を表すデータを作らなくちゃならない。それからようやく、コンピュータのデータ機能を使って、単純化したデータを通じて情報を管理できるようになるんだ」。
- いつか誰かが、情報を直接扱えるコンピュータを発明するかもしれない。しかし私は、それまで手をこまねいて待つつもりはない。そのようなコンピュータを発明するには、人工知能や自然言語処理といった、コンピュータ科学でも最も難しい問題を解決しなければならないからだ。その間、私たちは今のこの、頭は悪いが力だけは強い獣を利用しなければならない（本当にそんな新しいマシンが発明されたなら、それを作った人は、「コンピュータ」などという機械的な名前ではない、もっと新しい名前を付けなければならないかもしれない）。

次に続く2つのChapterでは、CMSが考慮しなければならないコンテンツの基本的な性質、「フォーマット」と「構造」について説明する。