

◆ The New York Times Information Bank

# 뉴욕타임즈 情報銀行

John Rothman 著  
李 炳 穆 譯  
(延世大學校 圖書館學科)

譯者註. 이 記事는 John Rothman氏가 'The New York Times Information Bank'란 標題로 *Special Libraries*, vol. 63 (April 1972), p. 111—115에 寄稿한 것을 *Special Libraries*와 筆者의 翻譯 許可를 얻어 우리 말로 옮긴 것이다. Rothman氏는 뉴욕에 있는 「뉴욕타임즈」社의 도서관 및 정보업무담당 도서관장으로 在職中에 있다.

Acknowledgements. This paper is translated into Korean and reproduced "from *Special Libraries*, vol. 63 (March 1972)"—Reproduced with the permission of *Special Libraries*.

대부분의 다른 신문들과 마찬가지로, 「뉴욕타임즈」의 뉴스기자들과 논설위원들이 이용해온 기본 정보자료는 오림자료 콜렉션 (clipping morgue) 이다. 오림자료 콜렉션은 약 60년전부터 구성해 왔으며, 그 범위는 「타임즈」의 각종 版의에도 60~70종의 각종 간행물에서 기자들이나 논설위원들에게 배후정보 (background information) 로 필요하리라 생각되는 기사들을 선정해서 오린 자료들이 포함되어 있다.

그러한 오림자료는 매주 11,000점에서 14,000점에 이르고 있으며, 지금까지 모인 오림자료의 수는 20만 점에 달하고 있다. 물론 그러한 오림자료의 대부분은 중복된 자료들이다. 이는 한 사건이 3~4개의 각각 다른 주제명과 4~5개의 제목을 지녔을 경우 보통 오림하는 방식으로 해당 주제 파일에 각각 끼워넣기 위해 필요한 숫자만큼을 오림하기 때문이다. 오림자료에 대한 색인은 색인작업에 오랜 경험은 가졌지만 정규교육은 받지않은 직원들이 처리해 왔다. 그래서 때로는 그 작업이 소홀히 다루어졌으며, 업무성경상 요구되는 일관성이라든가 정확성이 결여되기도 하였다.

뿐만 아니라 그와 같은 거대한 작업을 운영하는데 있어서는 색인가의 작업을 조정할 길이 전연 없었다. 또 각 오림자료에는 적절한 식별 표시를 하고, 자료를 배열해야 하는데 실제 오림자료가 제자리에 옮겨 배열되었는지 이를 확인할 길이 없었다. 이를 확인하고자

한다면 오림자료를 우선 해당 파일의 제자리에 끼우기 위한 노력만큼이나 많은 작업시간이 소요될 것이다.

두번째 문제는 일단 오림자료가 파일에 배열된 다음이라도 그 오림자료가 계속 그 파일에 보관된다고 장담하지 못하는 점이다. 오림자료를 이용하는 사람들은 집에서 이용하거나 혹은 복사하기 위해 오림자료를 파일에서 뽑았다가는 다시 제자리에 잘 끼워넣지 않는다면 때는 파일 전체가 없어지기도 한다. 간혹, 시간에 쫓기는 기자는 파일을 하나하나 따로 조사하지 않고 두, 세개의 파일을 한꺼번에 그의 책상에 벌터놓고 조사를 한다. 조사가 끝난 다음에는 원래의 여러개 파일은 온데 간데 없어지고 한무더기의 뒤편이 된 파일만이 남게 된다.

보관장소도 항상 문제가 되어왔다. 그래서 오림자료가 정리되면 필수로 점점 파일은 뚝뚝하게 채워져 나갔고 보관상태는 나빠져 갔다. 그 결과 오림자료는 구김살 투성이가 되었으며, 헤어져 읽을 수 없을 정도까지 되었다.

끝으로, 설혹 오림자료가 끼워 있어야 할 제 폴더 (folder) 에 끼워 있고, 읽어 볼 수 있는 형태로 그리고 올바른 연도와 날자순으로 배열되어 있다 하더라도 오림자료 콜렉션을 유지하고 이용하는데는 기본적으로 해결해야할 문제점이 남아 있었다. 즉 각 파일은 오로지 하나의 주제로만 검색할 수 있기 때문에, 예컨대 조사자가 '캘리포니아에 있어서의 해병대의 비행기 추락사고' (Airplane crashes, California and Marine Corps) 라는 주제에 관심을 갖고 있다면 그는 비행기 해병대 혹은 캘리포니아란 주제 아래서 찾아야하기 때문에 그러한 주제명을 포괄하고 있는 각 폴더를 뒤지게 마련이다. 그 경우 그 주제명아래는 물론, 비행기 추락사고라면 비행기란 특정한 주제속에 비행기에 관한 다른 모든 일반자료들도 끼워있다.

이러한 원리는 시비의 여지가 없다. 특히 상세하게 색인한 기사파일 보다 문서파일 (document file) 을

다를 경우 그러한 원리가 항상 적용된다. 문서는 한곳에만 파일되어 있게 마련이고 따라서 한가지 접근 방법, 즉 한가지 索引語 (index term) 로만 검색할 수 있다. 그러나 여러개의 색인이 가운데 자기가 찾고자 하는 바로 그 자료를 찾기 위한 정확한 색인어를 찾아낼 수 있는 사람은 아무도 없다. 이것이 곧 오림자료 파일의 기본적인 결점인 동시에 「뉴욕타임즈」가 이를 시정해서 「정보은행」(Information Bank) 을 세우게 된 주된 이유이기도하다. 다음은 「정보은행」에서 채택하고 있는 은행 운영 방식에 대한 설명이다.

### 入力 節次

「뉴욕타임즈색인」(New York Times Index) 에 들어 갈 그리고 오림자료 컬렉션 (clipping morgue) 에 정리해 넣을 오림자료 가운데 한 부와 런 오버(run-over) 는 이를 한 장의 臺紙에 풀로 붙이고 매지 윗 부분에는 연도와 일련번호를 기록한다. 그 번호는 곧 99개의 프레임 (frame) 으로 된 마이크로피시 (microfiche) 의 프레임을 나타낸다. 환언하면 풀로 붙인 각 오림자료는 99개의 프레임으로 된 한 장의 마이크로피시 가운데 한 프레임이 된다. 이와같은 특수한 형태의 마이크로피시를 택하게 된 이유는 「뉴욕타임즈」의 일주일분 오림자료를 그러한 형태의 마이크로피시에 수록하기에 적합하기 때문이다. 다만 「일요판」은 4개가 필요하다.

풀로 붙인 오림자료 원판은 현재 「미국마이크로필름 회사」(Microfilming Corporation of America-MCA) 에 보내어 거기서 마이크로피시로 만든다. 그렇게해서 만든 마이크로피시는 索引家 즉 「뉴욕타임즈색인」작업에 종사하는 바로 그 사람들에게 보낸다.

색인가는 비디오 터미널 (video terminal) 에 앉아서 터미널에서 직접 그들이 색인할 각 오림자료에 담긴 내용의 초록 (abstract) 을 작성하고, 이와함께 日字一面數—欄數 표시와 마이크로피시 번호 (프레임에 기록되는 번호)와 그 오림자료 검색에 적절한 색인어를 결합한다. 그들은 「뉴욕타임즈」 자료뿐만 아니라 「뉴욕타임즈」 기사 작성에 이용하는 60~70개의 다른 간행물에 대해서도 동일한 작업을 한다. 索引課 (Index Department) 는 그와같은 비디오 터미널을 26개 가지고 있으며, 그 외에도 누구든지 이용할 수 있는 28개의 터미널이 신문사 여기저기에 설치되어 있다.

초록은 IBM 360/40 컴퓨터에 넣는다. 마이크로피시 원본은 50,000장의 마이크로피시를 보관할 수 있는 마이크로피시 보관장치 (microfiche storage device) 에 보관한다. 이 마이크로피시 보관장치는 500개의 카아트리지 (cartridge) 가 들어갈 수 있는 거대한 커라우

절 (carousel) 로 이루어져 있는데, 한 개의 카아트리지는 100장의 마이크로피시를 넣을 수 있다. 이 경우 50,000장의 마이크로피시는 40년치의 자료를 의미한다 즉 「타임즈」 일주일분 피시가 6장, 「일요판」이 4장, 기타간행물을 이용해서 만든 일주일분 피시가 15장, 그렇게 해서 일주일분 피시 합계가 25장이 된다. 따라서 50,000장은 곧 2,000주분에 해당하며, 이는 40년에 해당하는 분량의 자료를 뜻하는 것이다.

### 스크린에 나타난 情報

마이크로피시 보관 검색장치는 컴퓨터로 운영한다. 이용자가 터미널에서 그가 필요로하는 정보에 해당하는 초록을 훑어 본 다음, 그 초록의 원본인 오림자료를 보고자한다면 그는 컴퓨터 건반 (keyboard) 의 단추를 누른다. 단추는 보관장치의 커라우절 (carousel) 에 신호를 보내어 해당 카아트리지를 식별해내게 되고 다시 그 카아트리지로부터는 해당 마이크로피시를 뽑아내어 건반과 연결된 xy 시스템 (xy coordinated system) 의 화면 (일종의 영상인자판—platen) 에 이를 보여주게 된다. 화면에 나타난 마이크로피시 프레임은 4개의 짜맞춘 텔레비존 (built-in television) 카메라 렌즈앞에 놓이도록 장치되어 있다. 카메라가 사용중이 아닌 경우, 이와같이 하는데 소요되는 시간, 즉 이용자가 「피시를 보여주시오」하고 단추를 누른후 화면에 피시 영상이 나타나기까지는 약 11초밖에 걸리지 않는다. 화면에 피시 영상이 나타나면 그는 피시 프레임 전체를 다 볼 수 있다. 이때 프레임을 똑똑하게 전부 읽을 수는 없지만 제목은 쉽게 읽을 수가 있고 따라서 그가 읽고자하는 부분이 어딘가를 결정할 수 있다. 터미널 곁에는 또 다른 조그마한 단추가 있어서 이를 누르면 그가 보고자하는 부분을 좀 더 자세히 살펴볼 수 있다. 이때 그는 텔레비존 카메라 줌렌즈 (zoom lens) 의 손잡이를 돌려 똑똑하게 읽을 수 있을 때까지 원문을 확대할 수가 있다. 이와같이해서 화면에 나타난 오림자료는 欄의 넓이가 2~3개에 이르고 길이는 30~40줄에 이른다. 이러한 보관, 검색 방식은 가장 흥미있는 그리고 개선된 업무처리방식, 곧 「정보은행」운영방식의 일부본일듯하다. 즉 컴퓨터에 의해, 거대한 파일에서 기사 원문을 찾아내고, 다시 이를 화면에 나타내는 일이라든가 복잡한 개념조정기술 (concept coordination technique) 들은 그전의 업무처리방식에 비한다면 놀랄만한 개선인 것이다.

「정보은행」시스템 가운데 이제까지 설명한 부분은 물론 社內에서만 이용이 가능하다. 기계도 특정한 수상기에만 송신되는 즉 클로우즈드 서킷 텔레비존 시스템 (閉鎖回路式 텔레비존 시스템—closed-circuit

television system) 이며, 현재의 기술적 및 경제적인 여건에서는 최대 투수량이 2,000ft에 지나지 않는다. 그러므로 마이크로피치 방식은 社内에 용으로 그 이용을 제한할 수 밖에 없는 형편이다.

외부이용자들은 대신 「타임즈」로부터 초록을 받을 수가 있으며, 그들이 어떠한 터미널을 이용하던 상관없이 터미널의 스크린에 이를 재생시킬 수가 있다. 그들은 신문 정기구독외에, 원한다면 마이크로피치도 구독할 수 있다. (이 경우 마이크로피치는 「뉴욕타임즈」 자료에 한한다. 왜냐하면 「타임즈」 자료가 아닌 것은 다른 출판사에서 판권을 가지고 있기 때문이다. 그러나 앞으로는 다른 출판사와 협의해서 그러한 자료도 제공할 수 있게 되기를 바라고 있다.)

초록은 「타임즈」 자료는 물론 「비타임즈」도 공급하고 있다. 그래서 외부 이용자는 그러한 초록을 이용해서 「타임즈」의 피치를 쉽게 이용할 수 있다. 그러나 「타임즈」 본사에 설치된 것과는 달리 마이크로피치 독서/인쇄결합기계 (microfiche reader/printer) 를 이용하여야 한다. 환언하면 외부이용자는 컴퓨터방식이 아닌 수동식방식을 택하게 된다. 즉 이용자가 마이크로피치 독서기의 화면에 나타난 초록을 살펴 본 다음 해당 자료의 원문을 보고자한다면 그가 직접 가든지 혹은 다른 사람을 보내어 번호순대로 따로이 보관해둔 마이크로피치 파일에서 해당 피치를 찾아내야한다. 수동식이지만 하지만 번호순대로 배열되어 있기 때문에 찾는 데 불편은 없다. 우리는 출판사들로부터 그렇게 하도록 허가해 달라는 요청을 받고 있으며, 거의 대부분의 경우 이를 허가해주고 있다.

## 그 외의 特徵

이제까지 「정보은행」 시스템이 어떻게 운영되는가를 매우 간략하게 살펴 보았다. 이용자가 색인가나 시스템技士나 혹은 사서의 도움을 받지 않고서도 자료를 신속히 그리고 쉽게 찾을 수 있도록 여러가지 기술들이 자료를 검색하는 과정에 도입되었다. 이 시스템은 「타임즈」 기자와 논설위원들에게 우선적으로 그 이용이 소개되었다. 「타임즈」 직원들에게만 이용교육을 실시한 근본 의도는 그들이 항상 마감시간에 쫓기면서 일하고 있는 점과 그들이 이용해온 옛날 방식과 새로이 이용할 방식을 비교해 보기 위함이었다. 오립자료는 외부인사들에게 개방되지 않았기 때문에 「정보은행」과 「타임즈색인」 및 「타임즈마이크로필름」 이용과 비교해 보는데는 외부인사가 안성 맞춤이었다.

이 시스템에 소개된 한가지 특징은 語典(thesaurus)으로서, 시스템은 이를 이용하여 자동적으로 색인어를 찾아내고, 자동적으로 참조(cross reference)와 범

위주기(scope notes)를 나타내며, 자료를 찾고 있는 이용자에게 조언을 제공한다.

또 다른 한가지 특징은 색인어에 전혀 의존하지 않고 조사의 범위를 좁히거나 넓힐 수 있는 기술이다. 예컨대, 자료 조사에 있어서 날자는 매우 중요시 된다 어떤 사람이 '사고에 관한 자료를 원합니다' 하고 말할 때, 그가 원하기만 한다면 비행기사고에서부터 눈 설매차사고에 이르기까지 모든 사고에 관한 모든 자료를 찾아 볼 수 있다. 이 경우 그는 여러 색인어(혹은 주제명) 밑에서 찾을 필요가 없고, 다만 대략 어느 기간동안의 것을 원하고 있는지를 알고 있다면——가령 3월 15일부터 3월 29일까지 사이의 모든 사고라고 가정해 보자——그는 그가 찾고자하는 자료의 범위를 그와같이 좁힐 수 있으며, 이에 따라 그는 사고에 관한 그 많은 자료를 하나하나 뒤적이지 않아도 된다.

이용자는 또 자료의 출처로 그가 원하는 자료의 범위를 좁힐 수 있다. 이를테면 「뉴욕타임즈」 자료 가운데 AP나 UPI 혹은 Reuters로부터 온 모든 자료를 각각 원한다면 그렇게 출처에 의해서도 찾을 수 있다. 그는 기사의 副題目이나 자료의 형태로도 범위를 좁힐 수 있다. 예컨대, 오립자료중에서 해답을 찾아내야 할 조사 가운데는 '정병소집에 대하여 편집자에게 보내는 편지'와 같은 조사가 있고, 이는 옛 시스템으로서는 조사하기가 매우 힘이 들었다. 즉 오립자료 컬렉션은 '소집'에 관한 자료가 담긴 폴더를 제공할 수 있기는 하지만, '편집자에게 보내는 편지'(letters to editor)는 소집에 관한 다른 자료와 함께 뒤섞여 있기 마련이었다. 따라서 옛 시스템으로서는 그와같은 자료를 찾기가 매우 힘이 들었다. 그러나 새로운 시스템에서는 간단히 '소집에 관한 자료를 원합니다. 단 편집자에게 "보내는 편지"만을 원합니다'라고 기계에 지시를 내릴 수가 있고, 따라서 그 자료만 볼 수 있게 된다. 물론 그는 여론조사, 뉴스해설, 사설등에서도 그와같은 방식으로 소집에 관한 자료를 뽑아낼 수 있다.

조사자는 또 그가 원하기만 한다면 특정한 삽도가 딸린 자료도 요청할 수 있다. 이를테면 신문사에서 작성한 월남전쟁지도도를 원한다고 가정해 보자. 그는 월남에 관해 지도가 들어 있는 자료요청을 기계에 지시할 수가 있고, 이에 따라 그는 지도가 딸린 자료를 얻을 수 있다.

이와같은 자료검색 방식은 검색을 좀 더 쉽게 그리고 어떤 의미에선 색인어와 상관없이 자료를 검색할 수 있게 장치된 방식 가운데 한, 두 가지 예에 지나지 않는다.

## 파일

정보은행을 운영하기 위한 기본 도구인 파일에 넣을 자료는 1971년 12월 1일부터 정리되기 시작하였다. 시

시스템이 완전히 가동하게 될 단계에 들어가면 업무시범도 실시할 계획으로 있다. 새로운 시스템의 데이터 베이스(data base)는 1972년 12월 1일부터 정리하기 시작한 「뉴욕타임즈」의 모든 자료를 포함하고 있다.

「타임즈색인」은 1968년부터 컴퓨터를 이용해서 만들어내고 있는데 이 「색인」작업은 이러한 새로운 시스템을 운영하는데 있어서 컴퓨터를 효과적으로 사용할 수 있을지를 결정하는 기회를 제공하였고, 결과 효과가 있는 것으로 판명되었다. 이 시스템은 또 자료를 기계로 읽을 수 있게 해주고 있다. 그러나 자료를 「정보은행」에 넣기 위해서는 약간의 편집작업이 필요하다. 그와같은 작업은 현재 순조롭게 진행중에 있으며, 1972년 가을까지는 1969, 1970, 1972년도분의 자료가 파일에 정리될 예정으로 있고, 그 다음은 1968년치를 다룰 예정으로 있다.

1972년말까지는 완전 정리된 파일——「타임즈」와「비타임즈」를 모두 포함——을 갖추게 될 것이고, 「타임즈」자료는 1968년까지 소급해서 정리하게 될 것이다. 다음 계획은 현재 오림자료 콜렉션에 들어 있는 오림자료들——그 가운데 얼마는 금세기초까지 거슬러 올라가는 자료도 있다——을 정리할 예정으로 있다. 이 작업은 5년 내지 10년이 걸리리라 보는데, 그렇게 시간이 많이 걸리는 이유는 기존파일(back file)의 보관상태가 불량하기 때문이다. 그러나 5년 이내——아마 1975/76년 쯤——에는 가장 중요한 자료의 대부분과 전체 오림자료 콜렉션 가운데 반 이상을 정리하게 될 것이다. 그 나머지를 정리하는 데는 약 5년이 더 걸릴 것

이다.

「타임즈」직원의 이용경험으로 보거나 이용자 조사를 통해서 볼 때, 社内 이용과 社外 이용을 포함해서 모든 질문 가운데 약 85%가 지난 5년 이내의 자료라는 것이 판명되었다. 그러므로 1975년경까지는 데이터 베이스가 대부분의 자료 요청을 충족시키게 될 것이다.

## 그 다음은?

앞으로는 어떤 자료를 축적하게 되는가? 만일 「정보은행」시스템이 성공적으로 운영된다면 「타임즈」는 데이터 베이스를 더 크게 확장해야되리라 본다. 즉 「타임즈정보은행」정보의 일부가 되고자 하는 간행물과 「타임즈」의 소프트웨어(software)를 이용하거나 혹은 「타임즈」와의 어떤 계약을 통하여 색인작업을 그들 스스로가 직접하고 있는, 출판사 혹은 소장자에 의하여 정리되는 간행물이 다수 포함될 것이다.

그 다음 단계는 사진자료를 정리하는 일이 될 것이다. 사진색인작업(picture indexing)은 기사색인 보다 훨씬 더 복잡하며, 현재의 형편으로는 그 작업이 원시적인 단계에 머물러 있다. 어떻게 사진자료를 색인하고 축적하고 검색하는 것이 적절한가 하는 것은 기사를 어떻게 색인하는 것이 적절한가 하는 문제보다도 알려진 바가 거의 없다. 「타임즈」는 상당히 큰 규모의 사진 콜렉션——약 200만점——을 가지고 있다. 사진자료 정리 시스템의 개발 문제는 사진자료 콜렉션을 많이 가지고 있는 다른 기관들과 협동해서 이를 처리하는 것이 가능할 것이다. (8면에 계속)

## 12면의 계속

과 同時, 特別 政府의 重要한 施策에 關한 事項에 對하여는 具體的인 장래의 예측등을 기술할 때 곧 政府의 公式見解로 하여 決定되고 있는 範圍의 것을 제외하고 事前에 그 要點에 對하여 閣議의 承認을 거치도록 한다.」 그리고 「白書類의 編輯 및 公表에 對하여는 責任體制를 確立할것, 政府刊行物 內에 正式書名(副題를 포함) 가운데 「白書」라고 하는 文字를 使用할 경우는 別表(本欄에서는 省略)에 記載하는 「白書類에 限한다. 더욱 白書類의 印刷, 發行에 對하여는 「政府刊行物의 普及強化에 對하여」(1956年 11月 2日 閣議承認에 定하여져 있으므로 特別 大藏省印刷局을 活用하는 것으로서 한다)라고 명시 하고 있다.

이와 같이 大量生産을 막기 爲해 統制와 規制를 加할 정도의 形편이나 그렇지 못한 우리 나라의 理由는 那邊에 있을까?

한편, 利用者의 경우에서 살펴보면, 政府刊行物은 多方面에 걸쳐 市中에 판매되지 않는 것이 있으므로 求得하는데 어려움이 있으나 적어도 白書의 경우에서는 대체로 市販되고 있으므로 利用者가 關心을 가지고 주의 깊게 관찰하면 입수하는데 별로 어려움이 없을 줄 안다. 그러나 白書는 다른 參考圖書와

달라서 특히 事前에 그 概要를 알지 못하면 活用하기가 어렵다. 예컨대 건국의 稅務署數와 位置를 알려면 「국세행정백서」를 보아야 하고, 世界主要國의 觀光선전 現況에 대한 것을 조사하려면 「觀光白書」가 무엇보다도 정확하고 仔細하다. 우리 民族의 所願이 國土統一인데 統一을 爲한 概要를 알려면 「통일백서」에서는 그 歷史的 背景에서 부터 政府의 努力, 國際關係에 이르기 까지 세밀히 기술되어 있고 獨逸의 分斷에서 부터 越南, 中國에 까지 多樣하게 소개하고 있다. 이와 같이 白書에는 그 題名과 관련된 모든 統計資料, 會議錄, 組織機構 改革 등 참고자료가 수없이 散在하고 있다. 또한 하나의 특징으로는 대개의 백서에는 附錄이 添付되어 있기 때문에 意外로 利用者에게 도움을 주고 있는 점이다.

이와 같이 白書는 Reference tool의 役割뿐만 아니라 所管事務를 把握하는데 貴重한 資料이면서 前項에서 言及한 바와 같이 어떤 白書가 刊行되었는지 어디에서 어떻게 구하여야 할지 일한 국민은 잘 모르고 있다. 하루속히 政府는 發行에만 意義를 두지 말고 活用하는데 있어 最善의 努力을 경주하여야 할 것이며 文化의 一翼을 맡은 全圖書館人도 이를 紹介하고 斡旋하는데 인색지 말아야 할 것이다.