

〔基調講演〕アーカイブの未来について

円 城 塔

アーカイブの未来についてということで、お話しさせていただきます。僕は主に SNS で電子データについて愚痴ったりしていることが多いですが、多分、その辺りが今日呼んでいただけた理由なのではないかと思っています。

アーカイブなので、何かを蒐集する。基本的に何かをするにはアーカイブが必要なんですっていうか、底本がないと基本的に議論が進まないはずなので、底本を定めるというのは大事ですね。ただ、蒐集されるものは作っている側の視点からすると、実際問題として集められているものというのはごく一部である。わりとアーカイブされない。わざわざ見せない作業というのは、結構な量として出てきます。出版社に残される作業記録とか、そういうものもあります。こういうふう《風》に恣意的な改変というのはすごく多いです。対談原稿に《などは顕著です》。

そもそも対談原稿とか鼎談原稿とかいうのは、しゃべったものをそのまま起こせるようにしゃべれる人はいないわけです。そういうふう《風》にしゃべれるのは、ほぼ浅田彰《さん》とかなので、他の参加者は適当にしゃべったのを、それを起こす人が頑張ってストーリーにするというので、その頑張る人の趣味とか能力とかによって結構変わる。

その後、恣意的な改編もかなり入ります。おもに誌（紙）面《の制限》によって。とくに新聞とかでそうですが、字数が足りないので削る。削ると論旨が

がらっと変わるなどのこともたくさん起こるわけです。ただ、それがアーカイブの資料として残って、後にその人の人柄なり作風なりを考えるよすがになるというのが、現場の立場的なものから見ると、起こっていることです。アーカイブは一応あるんですが、蒐集される側からすると、結構、好きなものを突っ込めるよみみたいな話でした。

一応、アーカイブには通常、書籍、雑誌、原稿などがあって研究の基盤になるもので、基本的にここで対象、底本を定めてから始まり《はじまり》ます。書籍は分かりやすい。雑誌に掲載されたが書籍化されなかったもの、国会図書館に行こう。雑誌掲載の対談、鼎談なども、国会図書館に行こう。パブリックなほう《方》はそこそこフィックスされて固いものですが、《次となると》遺品ですね。みんなやることがない《なくなる》と遺品を探し始める《はじめる》。当人が持っていた書籍とかですね。ラフカディオ・ハーンの本棚にはこんなものがあった。書簡ですね。あと作業用の資料、考古学的資料、住んでいた家とかになりますね。必ずしもテキストとは限らない。非常にプライベートなところ。この辺りが大体、研究対象とされるものですね。フィックスしたいようなしたくないような。

これはある程度、物なんです。紙、家、木簡などあるわけですが。とはいえさまざまなものが電子化しつつ、してない、してる、してる。基本的に作品を作るところと出力されるところも、もうほぼ電子化、機械化され終わりました。70代とかだと手で小説を書いている方とか、50代、60代だと手《紙》でマンガを書い《描い》ている方とか、まだいらっしやいますが。ただ、もうマンガ家さんはほぼ電子化が終わりましたね、書←《描く》のにおいても。後

でちょっと出るかもしれないですが、マンガ家さんは、もはや生成 AI も入っているんで、背景など書いていられないっていう、背景は生成 AI でよろしいのではっていうところを、生成 AI に対する是非はいろいろありますが。言ってももう入っちゃってるし、絵を描く側としても、もう学習されるのは前提で描きます、みたいなところまで、マンガのほう《方》はいつてます《いるわけです》。マンガの出力のほう《方》も電子化、もう売り上げの半分は電子に移行してると言われ、編集さんも別に電子で見て、ここ修正してくださいって電子で書いたものの版管理でいいので、実はマンガのほう《方》がバージョン管理できています《たりする》。そしてマンガは、最悪、ページを重ねればどこが変わったか分かるので、差異が取りやすいんです。コミックのほう《方》が電子データ的にはでかいんですが、実は、差異とかは見やすいです。

小説のほう《方》はなんかぐちゃぐちゃしています。原稿は大抵、「.docx」が来ますね。Microsoft の勝利ですが。わりと最近の学生さんは「.docx」と「.txt」の区別がついていないことがあってショックなんです。変換ってどうするんですか—《と訊かれて、》うーんって《なったり》。版下は大抵、InDesign とかです。印刷所が最終的に頑張るものですね。連絡は電子メール等なので—《、》僕はないですけど、LINE とかのこともあるらしく、書簡が消失しています。なので、遺品等の考古学的資料には、だんだんアクセスできなくなりつつありますね。書簡集がない。誰がどう頼んだのかもよく分からない。何にも残らないから交友関係も分からない、みたいな。実際《のところで》、交友関係—《は》ジャンルと関係ない《ん》ですからね《けど、そこは見えなくなる》。

~~これが~~《それで》、電子化の~~やつは~~、《話が》紙とか《の物質》と違うのは、たとえ電子データが残されたところで、そのデータが真正なものである保証がどこにっていうのが、今、兵庫県でたくさん起こっていることです。あれ、紙だったらまだよかったんだけど、そのフォルダの名前がクーデターだった~~か~~
~~な~~《のでクーデターである》、みたいな。そういう、その写真は正しいの、みたいなことは、電子的にはもうほぼ判別しようがない~~ので~~。でも、アーカイブとか大丈夫なの、みたいな話で《も》ある~~な~~《わけ》です。

今《は》、紙と電子の併存期なので、ちょっと具体的な話としては、目の前に電子書籍がありました。これはアーカイブされる対象、アーカイブされたと思っていいの、みたいなことなんです~~が~~。現状の電子書籍の版は、多分、多くの場合、紙の何かの版に依拠しています。電子書籍の一番最後の所まで送ってもらくと、これは紙版の何とかに依拠して電子書籍化しましたっていうのが、多分、最後に書いてある。最初の部分に~~「~~《「》ちょっと改変したかも~~「~~
《》》みたいなことが書いてあって、《「》ちょっと改変したかも《》》は許されるんだっけ、みたいなことですね。それは何だろう。紙版の版が変わったときの《電子版の》扱いについては不透明です。電子版において改変箇所があることが明記されている。ちょっと変えましたって書いてある。ちょっと変えましたはどこにっていうのは、分かりません。それは多分、編集者も知らなくて、書き手も知りません。~~なぜなら~~《少なくとも》僕《は》、電子版のグラをもらって赤を入れたことはない~~から~~です~~ね~~。それは多分、印刷所の人しか知らない。現状、何かの電子版を底本として扱うことは非常に困難です。同一性が怪しいし、ページの引用とか~~が~~《を》、どうやってやるの、みたいな話で。も

うアリストテレスみたいにやるしかないですよ。何《という著作の何》行目、みたいな感じでいくしかないんですが。その辺のコンセンサスは全然ないです。およびデータもその辺《そういう形で使いやすくなって》は《い》ない。

小説、アーカイブ、電子とは何ぞやみたいなときで、一つの立場としては、テキストデータだけに集中させてくださいというのは、僕の立場としてはわりとそうです。レイアウトとか知らんと。僕は、テキストデータを保持すれば一応、自作だというような立場ではいますが、当然、違う人もいます。それは京極夏彦さんのように InDesign で書き、ページは絶対にまたがない、みたいな。ページをまたがないのはいいんだけど、電子書籍で 1 ページ《の行数》を変えたらどうするの、みたいな話で。1 行の文字数を変えたとき、またぎますよねみたいなことなので。京極夏彦さんは多分、京極夏彦 AI とかを電子書籍リーダーに入れて、サイズを変えると自動的にレイアウトし直してくれるってやるしかないんですが。なかなか難しい。

ただ、コード、世の中にたくさんあるコードは、~~それもコード~~は基本的に人間が書くテキストなので、コードのほう《方》のアーカイブ技術はそこそこ進んでいます。小説は実行系の必要がない。もしくは人間が実行系のコードであるという見方はできる。でも、そこを考えると電子書籍にも実行系はあるわけですね。電子書籍、PC であつたり Kindle であつたり。そいつの中でなんかのプログラムがうごめいていて、そいつがテキストデータと呼ばれている、なんか分からない 01 ビット列を配列し直して表示してくれるというような、何かの種類のマークアップに対するレイアウトエンジンが存在するという意味で。

小説、普通はバージョンの管理はする《はずな》んです。コードなので。テキストなので。ただ、現状、全くできていません。

小説のどの辺がデジタル化できてないかというと、出版社です。書き手と印刷所はほぼデジタル化されている。みんなマンガをクリスタで書いている。そして電子書籍の印刷所。ただ、小説の、一番大きなのは、校閲さん《で》、どうしても画面では校閲できない問題とかがあって。校閲記号が書けないとか《、特化された視覚体験とズレる》ということがあって。ただ、そうすると、デジタルデータで入稿されてきたものが一回、レイアウトを組んで、紙にプリントアウトされて、それを校閲さんが校閲記号を入れて編集さんに戻す。編集さんはその校閲記号を見ながらゲラに赤ペンで書き入れ、それをスキャンして PDF で送り返して~~くれる~~《くる》。PDF ~~というか~~《自体もアレですが》、印刷してペン入れ、ペンはまだ完全にアナログなので、そこでアナログデータ化~~する~~~~ために~~《してしまっただけで戻ってこないの》、差分を取るなどが非常に困難~~です~~《になります》。デジタルデータ《を》扱ってる方はご存じだと思いますが、PDF ~~も~~《は》、テキストデータとしてはわりと死んで《い》ます~~ね~~。Adobe 自体がもともとレイアウト用のソフトウェア~~な~~《の専門だった》ので。あいつは位置情報しか持っていないので、文章の構造は持っていません。入れようと思えば今《だと》、入りますが。

このフローは、《別に》いい《ん》ですけど、絶対に間違えるんです。デジタルから組んで紙にするときはいいですけど、紙に書いてデジタルに戻すときに、絶対に間違える。プログラミング業界だとそんなことは許されないの、常時差分を取りながらどこが変わったか確認しながらいくんですが、一回、ア

ナログデータ化するので差分は取れなくなるんですね。ということについて疑問は感じないのですか、というインタビューを僕は繰り返していますが、感じていませんって返事しか返ってこず、取りたくないですかって聞くと《訊く》んですが、意味が分からないとしか言われませんか。

小説が何とかそれでもいいのは、できちゃうからです。原稿用紙 200～300 枚だと、頑張れば全部、見れるっていう。徹夜すると何とかなる、みたいな。あのサイズがいいところにいて、わりと一晚二晩、作業すると何とかなるぞっていうのが、デジタル化の進まないところです。デジタル化すれば楽なのに。

もうちょっと範囲を広げて、今、小説を書きたいとか言ってもクリエーティブ《クリエーティブ》ライティングとかに来る人の《と》話を聞いて《したあとで》、正直何をしたいっていう《訊く》と《、》ノベルゲーが作りたいですとかいう人のほう《方》が増えていて。卒論もノベルゲーで書きたいとか言うんですけど。でも、ノベルゲーの底本ってなんですか、みたいな話はありますね。

特にオンラインゲーム《です》。オンラインゲーム、ソーシャルゲームとかになると、実行系が維持されるかは《が》、そもそも心もとないです。これ、ゲーム批評がいまいちうまく機能しない一因でもあるんですが。《、》昔のゲームはエミュレーターで動かせばいいじゃんみたいな話は、まああり《ます》。『パックマン』とか動くかもしれないですが。でも、オンラインゲームは基本的にシステムのスナップショットとしてしか存在しないので、MMO とか、人がいっぱいいるオンラインゲームですね、いったん停止して再起動できるか

どうかは不明ですね。10 年前に止めたオンラインゲームを再起動しますとか言って、動かないと思うんだよな。オンラインでも、単独で遊ぶゲームは、起動ぐらいはできるのではないかな。このあいだ、なんでしたっけ←『メギド』が終わったんですよね。《あれはオフライン版に移行するそうですが。》そういうの《は》ありますが。ただ、任意のバージョンを実行するのは結構、難しい。『Minecraft』は今のところできています。あれ中身←《が》全部見えるからなんですけど。中身、全部見えるから、MOD はこのバージョンじゃないと動かないとかいうような生態系が、なぜかできてますね。でも、逆に言うと、あれぐらいしかできないかもしれない。だからもう『FGO』で卒論書きたいです。奈須きのこファンですみたいな人が現れたときに、じゃあ、おまえ《お前》の《扱う》底本は何なんだよっていうときに、自分が~~ブレ~~←《ブレイ》したスクリーンショット全部なのかみたいな話になって、もう《わけが》分からないわけですね。こうなると、アーカイブって何なんだろうみたいな話~~があり~~《になり》ます。

でもアーカイブが整備できれば研究ができる。映像、音楽作品は意外と最終データが分かり~~やすいので~~《やすくて》、公開時のフィルム／データがある程度、固定されるので触りやすい。映画館をまわっているフィルムが底本であるという考え方があるし、制作側が保持しているのもそこそこある。当然、捨てられていくデータとか猛烈にあるはずなんですけど、これって~~の~~《いう候補》が《一応》出ますね。

小説、コミック等、紙メディアと併存しているものは紙の底本が一応ある。ただ、紙にならないものも増えています~~ね~~。電子版しか出ないマンガとかいつ

ばいあるので。電子版しか出《出版され》ない小説~~で~~《は》あんまりないんですけど、電子版しか出ないマンガが、今~~すごく~~《すごく》増えてますね。そのデータがどこで管理されているのかはいまひとつ不明です。不明って、単に僕が知らないだけなんです。Kindle が持ってるデータと、BOOK☆WALKER が持っているデータとかは同一なのとか分からないですね。BOOK☆WALKER って KADOKAWA が持っているんですね。《いて、》KADOKAWA が持っているのを他のプラットフォームが使っているんですよ。なんかよく分かんないぞみたいな状況を、この間、知りました。KADOKAWA のライトノベル系忘年会、忘年会じゃない、なんかの会に忍び込んでそれを聞いてきたんですが。

ゲーム等はアーカイブがほぼないので、いろいろ起こりがちです。そうはいうけど、じゃあ~~何な~~《どうすればいい》んだよって言われるときに、《とりあえずは》原稿の管理《ですか》。僕は、もう GitHub で動かしている。これです。

https://github.com/EnJoeToh/stories_2000

これは僕の GitHub の公開しているページですが、ページですがって言っても、どこかに載った原稿じゃなくて、他の原稿も最近~~これ、かなり~~《こんな感じで》管理するように~~な~~《し》ていますが、これは 2000 字くらいの短編を 100 個、置いてある。何のために、みたいな話なんです。機械学習とかに使って。後《時間があれば、あと》で出す《学習結果の例まで.....いけない》か。機械学習に便利に使える小説データ《が》ないなっていうのが、一番最初

の動機ですが、この辺は《に、》 こういうデータがあるわけです。ヒストリーで更新が見られるようになっていて、~~ブラウザ~~ヒストリーで《をブラウザできます。こことか押すと》。《.....》これ変えてないやつだね。変えてると、ここにログが見えるんです。~~次の紙に書いてあるんですけど。~~《ええと、》 29番に行き、ヒストリーを見ると、~~これ~~《.....ここ》を広げると、こんななっていて。これ《が》誤字修正したやつですが。差分とかを見れる~~んですけど~~。どこで何を変えたかをコメント入りで差分が見える。Twitter で誰かが誤字を教えてくださいましたので、誰が教えてくれたかは、あの方が教えてくれたんですが。こういうのですね。え、~~変わる~~《どこも違わ》ないじゃないと思うんですけど、「シバ刈り」、そうだよな、おじいさんが刈りに行くシバは、木の~~ほう~~《方》の「柴」であって、草刈りに行くわけじゃないんだよねっていうシバ刈りの字が違う。ていうのが、全部バージョン管理ができる。だから、こういうので履歴を管理しようっていうことです。ていうのを基本的にこの 4《4》年ぐらいで書いてるもの《で》は《やっていて》、~~僕はその~~データはこういう感じで持っています。GitHub のアカウントを取っていただくと、プライベートアカウントで編集さんに入ってもらって、ここで共有すればいいじゃんって話《なん》ですが、《「うん《》》って言った人が一人もいません。《「《》じゃあください《》》とか言う人もいないです。なんでなのかは、もはや僕に分かることではない。僕にできることはしたみたいなことではあるので、あれなのですが。というので、《ごくふつうに》管理をすればいいんじゃないか。僕は基本的にここまでできれば、小説の底本の管理としては別に悪くないのではないかと《と思っています》。やりとりで返ってきたものをここで入れる。ただ、自

分の修正指示が反映されているかをどう確かめるかとかは《残っていますが》。でも、この~~ほう~~《方》がやりやすいです。ここまでできれば、書き手の側としては終わり。とりあえず、小説のアーカイブとしてはいけるのではないかと思っている《います》。

というので、履歴管理はできます。小説だけじゃなくて法律とかもあります。変更の履歴を見れるのは、ワシントン D.C.はやってますね。コロンビア特別区の法律、結構でかい範囲でやっていて。

<https://github.com/DCCouncil/dc-law>

~~一応~~、DC Council のページを見に行くと、GitHub に置いてあるんですが、見に行くと、一応、GitHub の~~ほう~~《方》が本体だって書いてあって、え、そうなの、みたいな。ただ、XML で書いてあって読みにくいんですけど。一応そっち側の~~ほう~~《方》を底本にして、ウェブページの~~ほう~~《方》に上げているらしい。他にも、フランスとかでも法制定のときに、法のどの案をどういう~~ふう~~《風》に変えるっていうのを全部、追えるようにしているとかいうのは、一応やっている人はいます。アーカイブとして正しい使い方である。日本でも、日本国憲法とかをそれでバージョン管理しようとかしている人がいますが、みんな日本国憲法に興味がないので、全然、盛り上がらないらしいんですけど、やっている人はいます。

アーカイブとして整理されたデータであるなら、むしろ機械学習の対象になるはずですね。僕が GitHub に置いて公開しているのは、機械学習の対象にしてくださいと言ってや~~ってるんです~~《ています》。足りないんですけど、本

当はあれ 1000 ぐらい書きたかったんですが、~~そんなことやってる~~時間がなくて、今、100 で止まってるんですけど。あれ 1000 ぐらいあるとそこそこいくんじゃないかなと思ってはいます。そうすると、研究の対象というより、どっちかっていうと機械学習のデータになって、模倣がいろいろ問題になってい《き》ます。今、《画像系だと、》もう 1 枚か 2 枚、~~わりと~~原画があると機械学習されてしまうので、~~特に画の人はそうですね~~《しまったりします》。新しいキャラクターが出るとすぐさま学習されて、服を脱がされて販売されるっていう。もうあのループ何とかしろよって感じなんですけど。~~それとか~~《そういうことが》起こっています。ただ、どちらかという自作を学習させてアシスタントとして使う方向性のほう《方》が、今後、残りそうではあります。さっきも言いました、コミックの背景であるとか、いろんなキャラクターのポーズであるとかも、言うとか大炎上するので言わないけど、使っている人はわりと多いはず《です》。学習されることは織り込み済みな上で作るとかいう人のほう《方》が増えてい《ます》《いるはずですが》。一応、ナイトシェードを作るとか使うとか、ウォーターマークを入れるとか、サイズを細かく切り刻むとか、いろんな《対策》があるんですが、結局、乗り越えられるんで、~~最終的~~《られてしまうこと》には~~そこ~~《なる。もともと、こう》いうと《見方は、》諦めが早過ぎるかもしれない《んですが》。

小説のほう《方》は、学習に対して何か思ってる人は、学習禁止しますっていう人もいるんですけど。僕はわりと学習されてしまったほう《方》がありがたいという立場ではい《ます》《です》。なので本当は文芸誌とかに載ったやつデータも公開しちゃいたいんです。出版社が許さない《と》の《折り合い》で

できないんですが。とかいうのはあります。それはなぜかとかいうのを話している時間は、ちょっと後でいけばいきます。

データと作家性という話なんですが、結局、文学研究だと残されたデータからその人の「作家性」なりを復活させようとする、小説はそこそこの一人ですべて書いてるんで《わりといけます。ただ》、先ほども言いましたように《が》、インタビューや対談は結構あてになりません。わりとみんな自分がかっこいい方向に直すから。プライベートなデータは秘匿される傾向にありますね。《[]》

《 》コミックはいろいろですよ。とくに日本の場合、週刊連載とかもうめちゃくちゃですから。アシスタントの作業箇所はどこなのかというのと、作成効率など、もろもろいろんなものが絡んでくるので、どこまでが作家《個人の作業》なのか、目《以外》は描いているのか、みたいなこともいろいろあって、よく分からない。

映像作品になると、まあまあ参加人数が多いので、現場では経済的・技術的問題がそもそも大きな枠組みになる。小説であんまり経済的・技術的問題が大きな枠組みになって、それゆえに何か《を曲げる》ということは、ほぼないんですが。取材の効率とかありますけど。でも、コミックはかなり時間の~~ほう~~《方》に縛られてますし。映像作品は予算規模と、誰と知り合いなのかとか、いろいろありますね。

アニメーション作品の作家性は、監督とか脚本に対する評がスタンダードですが、別に、脚本に《が》書いたものがそのまま映像化されるなんてことはないわけです。特に日本のアニメーション~~であつたりする~~《の》場合。この頃のハリウッドは組合がうるさ過ぎて直せないらしいんですけど、直せないと映画

がつまらなくなるって話があって。じゃあどうするんだ、映像にできなきゃしかたないじゃん、などのこともあり、誰が修正したのか分からない箇所などもさまざま出てくるわけですね。

僕も『ゴジラ S.P.』の脚本というのをやりましたが、あそこはどうしてなんですかとか聞かれても、いや、そこ僕がやったところじゃないから分からないとかいうことはたくさんあり、そうすると、それを作家性とかに組み込まれると困るなっていう。そもそもゴジラのやつは、僕がなんで呼ばれたのかよく知らないですけど、SF っぽいことができる人っていうのと、ある程度、癖のある人を呼んでその癖を出したほう《方》がいいという癖役として呼ばれてるので、やり過ぎとみ言われても、僕は抑えたかったんですよ、どっちかっていうと《となる》。でも、「もっと円城塔っぽく書いてください」とか言われて、ああなってるん《わけ》ですよ。とかなので、作家性って言われるとちょっと困るなというところもあるんですが。でも、大体そんなもん《ですよ》。だから、虚淵玄さんとか大変ですよ。脚本を渡しただけなのに出来が全部、虚淵玄のせいになったりして。被害担当艦ですみたいなことになるんですけども。そこでカットの切り方とか、使われてる技法とかは、「空気感」とかふんわり流されますが、実はそっち側のほう《方》を見たほう《方》がいいのではとかいうのと、いろんなデータが残る《残ってる》んであれば、機械処理でその辺できそうなのにとかいうのと、そもそも制作サイドに訊いてみれば、裏事情をいくらでも教えてくれるのにみたいなこともあります。

映画評でも、レンズの明るさとか焦点距離とかはあまり気にされないですね。せいぜい手持ちカメラでやると、ああ、なんかぶれ《いい感じにブレ》ている

みたいな感じではありますけど《とこです》。なので、わりと映画評、アニメ評とかを、小説とかと同じ方向性の評価の軸で、昔は見て~~っていうふうに~~《いたわけです》。~~そこに行~~《自分が作品とし》て~~見~~《目の当たりにし》たものを何かする~~みたいなの~~と《いう形で》、作業過程と完全に切り分けられていたんで、~~それ~~《作品そのもの》が対象だったし、それに対して《ものを》言えばよかったんですが。今はわりとあらゆるものが見えるので、制作サイドに訊けばいいのにみたいなことってまあありますね。

ゲーム評は、歴史ものは特に主観になりやすいので、何とかゲーム史は大抵、炎上しますね。アーカイブがあやふやであることと、網羅性の低さがどうしてもあるので、独自史観が形成されがちなのと、あと、関係者が存命のことが多いので、関係者が SNS で文句を言ってくるっていう、それが違うっていうのを言ってくるのがあって。制作サイドに訊いてみればいいのに、みたいなことはあるわけです。《 》

《 》はて、小説の話だったのではってことですが、《そのあたりまで含めての「 》テキスト《 》」の話でもあります。

小説はそこそこの作業ではあって、編集さんとの関係はいろいろで、ジャンルにもよります。ライトノベル業界であると、いまだに編集さんのほう《方》が主体で、お話を組んでくるとか~~あるので~~《があります》。こういうお話を考えたんですけど書いてもらえませんか。それは自分で書けばいいのでは、みたいなことはジャンルによってはよく起こります。

出版という生態系に乗ってはいるんですが、これは今、激変期にあるので、もう《今後どうなっていくのか》分かりません。今日も文フリですが—《、》

もう文フリのほう《方》が実際、書き手に入る収入が増えた場合どうなるんですかみたいなことはあり。しかも《すると今度は》、文フリの同人誌のアーカイブとはなんであろうみたいなことになるんですが。《┐》

《音楽などで顕著のように、》作家性はまあまあ見出せたりし《強くなっ》て《いるような気がし》ます。《が、》公開されてるデータがむしろ速やかに消えつつあります。昔よりも。みんな、だってデジタルデータで、死ぬとiPhoneに誰もアクセスできなくなるので。ん？っていうことですね。

葬式とかも簡素化してゐるので《てきて》、最近、有名な人が死んでも、みんな身内で執り行いましたので、とかになって、そこで関係者が顔を合わせることもとかもなくなって。むしろデジタル化によっていろんなことが風化してる感じはあります。そ《と》いうので、ふとわれに返ると、でも、書簡集とかはわりと残そうとして残ったんですよね。正岡子規は自分が出す手紙も書き写してたわけじゃないですか。ちゃんと控えを取っておいた。ただ、正岡子規がおかしかったとかいう話ではなく、昔からの、そこそこ習慣としてあったわけですね。手紙は控えておきましょう《、という》。カーボン紙とかそのためにあったわけですね。複写を取るためにあったわけですから。習慣としてそこそこあったものです。日記を残してる人もまだそこそこ多い《はず》ですね。特に政治家とかに多いじゃないですか。細川家とかきつと残《残し》てるに違いない。なのでわりと残そうとして残った人のものが残っている。漱石とか、だって手紙焼いてたわけじゃないですか、庭で—《、》とかなので、実はその辺はあんまり変わらず残っているのではという気は、このお話をいただいてからちょっと考えそ《ました》。単にお調子者がデジタルに移行して、勝手に ✕

《X》とかで騒いでるだけなのでは、とかいう気はちょっとしまして。書簡と e-mail の違いっていうのは、多分あります。気軽さにおいて。~~そっち側も~~《メールというメディアに文体が》引っ張られてる《れるという》のはあるんですが《あります》。すると結構、作り手の意識次第なのでは、という気が最近してきました。歴史を~~クリエイト~~《クリエイト》しようという意志次第なのでは、と。減退してるのは、歴史を残そうとする意志なのでは、とかいうような気が、僕はしてきていて。最近、メールのやりとりと~~とか~~をするとき、もっと手紙っぽい文章で書くのはどうか、という気がしてい~~て~~《ます》。むしろ書簡集を自ら~~クリエイト~~《クリエイト》していく。でも昔の人もそうだったんだよな~~って~~いう気がわりとして。むしろ欲しいのは秘密の日記サービスであるとか、公開されることが前提のテキストのやりとりサービスなのかと思いつつ。でも新しいサービスっていうか、死後公開されるっていうステータスがデータに付加されていれば、まあいいのかと。

僕の原稿は **GitHub** 上にありますが、僕が死ぬと、多分、アクセスできなくなるんですよね。せつかくあるのに。それは遺書に書いておくとか、遺書にパスワードを書いておくとかいうことになると思うんですが。死によって失われるものは、無論、多いわけですが、その際に残されるものに「データ」《が登場した》。データを残す。そのデータが残ったことにより、これまでより「精度が高い」状態で故人が「再現」されることが起こり得るかもしれないんですが、ただ、今ここまでの話の流れでいくと、明らかに自分が残そうとしてるものを見せるので、より欺瞞の程度が上がっているだけであって、精度が高くなっているわけではないかもしれない。そもそもそれは「故人」ではない

ですし、「再現」でもないですね。美空ひばりが NHK で歌っても、それは別に精度の高くなった美空ひばりが新曲を歌ってるわけではないです。当時の状態のまま、ゾンビのように歌わせたらこうなるんじゃない、みたいなもので、それを故人と見なすのはむしろ問題がある。とかなんですが。

結局、フェイク、フィクション、トゥルースみたいなものの境界が、わりとデジタルデータを《では》保証できないので、それを何らかの形でひも付けながら残して、その人はどうだったっていうときに、やっぱり架空の度合いがすごく上がっているのではないかっていうのが、僕の今、思っていることです。なので、研究としてはどうするんでしょうね、みたいなところは、かなり大きな問題として思っています。

話はちょっとずれるんですが、機械学習みたいな話で、自分のデータをどうしたいかっていうところで、僕は学習してほしいって話をしたんですが、『文字渦』という小説を書いて、日本語と遊んでいるうちに日本語がよく分かん《分かん》なくなってきて。日本語ってなんかよく分からない~~夫和~~《、やまと》言葉と中国語の、超中国語方言の混淆なんじゃないの、って気持ちになって。あと、仏教用語が入る。仏教用語に規定されたローカル中国語に~~夫和~~《、やまと》言葉が交ざっている、変な書記体系であるっていう実感がすごく増えてしまい。《、》結局、日本の思想は仏教用語のほう《方》から入ってきて、ここからしか語ることはできなかったのではないかと、みたいな疑問などが兆し。《、》それは、日本語というもののの中に仏教要素が勝手にインストールされたからですね《ているように見える》。勝手にインストールされたので、思考の

枠組みまではいかないですが、その語彙を使わなければ何か説明できないという形で、近代、現代日本語が成立している《と感じたわけです》。

というような中で《そうして》、~~あと~~海外から日本文学をデータとして扱うみたいな感じの計量文献学みたいなことをやるときに参照されるのが《はま

ず》、青空文庫~~である~~《ですね》。とはいえ、それはもうほぼ 100 年前の日本語なので、《使いたくなるのは》分かるけどっていうときに、もう俺の読めみたいな気持ちは多少~~ある~~《起こる》わけですね。さらに、日本語はどうせ変化していくものなので、そのときに仏教代わりにはと言わないですけど、仏教要素的なものの代わりに、わりと理数的な語彙なり何なりを差し込みたいというのが、僕の中の欲求としてあってですね。そうすると、そういうの《テキスト》を突っ込み、資料として与えておいて学習させると、そちら側の~~ほう~~《方》が日本語~~で~~《として》学習された内容として ChatGPT とか~~で~~《から》出てくるわけですね。なので、そっち側に、機械の~~ほう~~《方》にこちらからデータを与えて、むしろ日本語を侵食するチャンスであるっていう~~ふう~~《風》なところは、僕は今、思っているところです。

なので、アーカイブの話とはちょっとずれるんですけども。ただ、今後の日本語をわりと GPT なり何らかの生成 AI が決めていく可能性は高い~~ので、~~
~~高い~~と思われるので、積極的に関与していかないと、日本語は 100 年後、青空文庫文章に戻るんじゃないの、と。特に小説の学習データってあまりないので。むしろ小説家はそっちの、自分が作りたい日本語で書いて、それを学習させて完成させた~~ほう~~《方》がいいんじゃないの、と~~か~~いうのがわりと僕の思っていることです。

というのは、実際、画像生成 AI のアニメ《絵》は大体、日本で書か《描か》れたものが侵食して**いるので**《いますね》。あれは大変ひどいことであると同時に、侵食できるんだって**いう見本にもなっている**ので。なんでみんなそんなに無限に『エヴァンゲリオン』の画を作り続けるの、ってことですね。あれが《は、》「人新世」を表象するデータとして『エヴァンゲリオン』があるみたいな、よく分からないことになっていると僕は思うんですが—《、》アニメの覇権はその辺のものが握っているのではないかなと思い、アーカイブされたものと《同時に》生成 AI とかが**あ**《存在してい》て、もはやスナップショットを撮る意味《撮って瞬間を保存する方法》が**どうなる**《あるの》かも分からない**中で**《まま》、ぐるぐる回っていく中で、実製作側の**ほう**《方》は《ともかくも》やっていくの**だらうな**《だ、》というのが、今、思っていることです。**その辺**《AI まわりの規制は》、国にもうちょっと頑張っ**てほしい**ですね。

生成 AI だって、日本の同人誌とかアニメーションとかを勝手にあげる海外サイトが学習されてるわけじゃないですか。あれにみんな猛烈にタグを付けていたから学習でき**てるの**《ているわけ》で、駄目なんじゃないって。駄目ですよ。とかいう辺りはどうするの**か知らんとか**《かしらん》いうところで。〔 〕
《 》基調講演**な**《として適切だった》のか《は》分からないですが、取りあえず、ここまでとしていただければ《と思います》。