

Annex 2: Risk Based Verification

Annex 2: Risk Based Verification	1
Summary	2
Introduction	3
RBV is being used...	5
RBV has, in practice, significantly changed how 'High Risk' applicants and claimants are treated.	6
Background to RBV, and its interactions with the benefit system	9
Credit reference agencies	11
What does the RBV process look and feel like to a claimant?	13
What have we learned that can help people?	14
Is RBV making problems worse?	18
Testing RBV against 'value for money'	19
Algorithmic assumptions about 'fraud' and 'risk'	21
Is RBV necessary and proportionate?	22
Intellectual Property, Algorithms, and Artificial Intelligence	25

Contact: coordinator@medConfidential.org

Summary

The rules applying to Risk Based Verification (RBV) are very clear: the risk group to which a person is assigned by the (commercial) algorithm **cannot be downgraded**;¹ it can be *upgraded*, for a whole host of reasons – and there is no process of appeal.

A raft of new measures and processes are therefore applied to the highest risk group. DWP's insistence that the use of RBV has to be "cost neutral"² has the direct result of targeting all resources saved (staff time, etc.) by *not* checking 55% of applicants as rigorously as before, entirely onto those deemed 'High Risk'. 75% of checking is targeted against the 20% of applicants deemed to be 'high risk', with the remaining 25% of resources targeted at the 25% deemed medium risk.³ Those measures targeted against the 'high risk' claimants could not be applied to everyone, simply due to the level of resourcing that would be required to do so.

The basic approaches of RBV, even if not referred to by that specific term, are in use across UC and its fraud process – and this is most clearly visible outside the perimeter of DWP itself in local council-run aspects of the benefits system, where it is done to DWP's design.

In general, when systems give out different information based on different inputs, their use can *always* be detected, and *usually* tested.

Every tool used in decision-making by a public body must produce something which can be acted upon – and those actions can be constrained by law. Equality Impact Assessments for RBV are not routinely produced or available, but the analysis audit functions that currently are could be meta-analysed at a national level to measure systemic effects that individual councils are each too small (or under-resourced) to do themselves.

As part of their annual audit obligations, councils have begun to perform a 'value for money' analysis of RBV in the administration of Housing Benefit and Council Tax Support. RBV is failing those value for money tests, even after 'encouragement' and incentives from central Government have been factored in, increasingly so as UC rolls out (with benefit passporting).

The arbitrary nature of RBV 'cuts holes' in the social safety net with no warning, explanation or recourse for those individuals who fall through them. And, thus far, the questions raised in this report remain unanswered by the State.

¹ e.g. paragraph 7.3, <https://buckinghamshire.moderngov.co.uk/Data/AVDC%20Cabinet/201709061830/Agenda/Housing%20Benefits%20-%20Risk%20based%20verification%20policy.pdf> - "**Cases cannot be downgraded at any time** by an assessment officer, they **can be increased** though with approval from a Team Leader or Senior Officer. Reasons for upgrading a case may include **previous fraud, previous late notification** of changes in circumstances, **or where there is good reason to doubt the veracity of information provided...**"

² page 3, paragraph 10: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/637068/g1-2016.pdf - "**The introduction of RBV by an LA should be cost neutral, i.e. better direction of resources towards high risk cases and not less resource being utilised.**"

³ Table on page 7: <https://democracy.west-norfolk.gov.uk/documents/s3445/Housing%20Benefit%20and%20CT%20Support%20Risk%20Based%20Verification.pdf>

Introduction

When treating people differently, ‘means testing’ is at one end of the acceptability spectrum, with ‘literacy tests’ – which have historically been used to, e.g. deny certain people the vote – at the other.

“Risk” is a subjective measure.

“Risk-based” approaches allow one to vary ‘levels of acceptability’ by using independently variable levels of discrimination. The choice to use risk-based processes facilitates doing different things to or for different people for some reason; while some of those reasons can be well-intended, even good, those reasons can also be discriminatory or unlawful.

While any individual decision may be entirely arbitrary or properly justified, *bias* is only measurable at a systemic or system-wide level. No individual being profiled is likely to be able to meaningfully do anything about it themselves (cf. police ‘stop and search’ policies, and practice).

So, for example, when people are asked for their paperwork when applying for Housing Benefit, a risk-based approach means that some of them will have more hoops to jump through:

Type of Evidence	Sub-category of evidence	Low Risk	Medium Risk	High Risk
Identity and S19	Identity	Originals or Photocopies/ Scanned copies	Originals or Photocopies/Scanned Copies	Originals required
	S19	Originals or Photocopies/ Scanned copies	Originals or Photocopies/Scanned Copies	Originals required
Residency/Rent	Private Tenants		Originals or Photocopies/Scanned Copies	Originals required

However, any risk assessment process that creates a *written* or *recorded* output can be tested by non-government entities – because such decisions are measurable.

‘Stop and search’ lost confidence, not because any individual decision was necessarily wrong, but because – once all of those decisions were seen in aggregate – different groups of people could see that different criteria had been applied (to them). And for any administrative system, decision-making within it can be measured based on what the citizen is asked for, following the precedents of the JCWI case,⁴ Dr Byrom’s measures,⁵ or similar.

⁴ <https://www.bailii.org/ew/cases/EWHC/Admin/2019/452.html>

⁵

<https://www.thelegaleducationfoundation.org/articles/the-legal-education-foundation-is-today-publishing-a-blueprint-for-digital-justice>

Following DWP's 2011 Housing Benefit and Council Tax Benefit Circular,⁶ one of the first councils to voluntarily apply Risk-Based Verification, Leeds City Council, stated in its 2012 report on RBV to its Corporate Governance and Audit Committee:⁷

As reduced verification is already practiced on claims administered in Jobcentres and the Pension Disability and Carers Service the DWP allowed local authorities to pilot risk-based verification in Housing and Council Tax benefit claims. It is the DWP intention for risk-based verification to be applied to Universal Credit claims.

The report further stated:⁸

The risk models used in the DWP research were developed over a two year period to accurately and consistently differentiate claims across three categories of low, medium and high risk and were subject to rigorous statistical evaluation.

This and further, more recent, evidence indicates that DWP continues to pursue an ongoing programme to develop a “fully automated risk analysis and intelligence system on fraud and error”,⁹ using machine learning, an increasing amount of data from credit reference agencies and other capabilities. What we can see around RBV at the local authority level is thus only the tip of a very large iceberg of automated decision-making.

The rules applying to RBV are very clear: the risk group to which a person is assigned by the (commercial) algorithm *cannot be downgraded*. It can be upgraded, for a whole host of reasons; There is no process of appeal.¹⁰

Given people's rights in regard of automated individual decision-making and profiling (i.e. the automated processing of personal data to evaluate certain things about an individual) it would appear that the use of any RBV algorithm over which – for the vast majority of cases – the assessment officer has limited or no discretion in some regards, and against which the claimant has no possibility of appeal, is a *prima facie* breach of Section 14 of Part 2, Chapter 2 of DPA 2018, i.e. Article 22 GDPR:

Q1: Did DWP inform local authorities of their legal obligations with regard to RBV under DPA 2018?

6

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/633018/s11-2011.pdf

⁷ paragraph 2.4, <https://democracy.leeds.gov.uk/documents/s65921/RBV%20report%20AG%20Committee%20March%202012.pdf>

⁸ paragraph 2.3, *ibid.*

⁹ page 61, <https://www.nao.org.uk/wp-content/uploads/2018/06/Rolling-out-Universal-Credit.pdf>

¹⁰ e.g. paragraph 7.3, <https://buckinghamshire.moderngov.co.uk/Data/AVDC%20Cabinet/201709061830/Agenda/Housing%20Benefits%20-%20Risk%20based%20verification%20policy.pdf> - “Cases **cannot be downgraded at any time** by an assessment officer, they **can be increased** though with approval from a Team Leader or Senior Officer. Reasons for upgrading a case may include **previous fraud, previous late notification of changes in circumstances, or where there is good reason to doubt the veracity of information provided...**”

- If not, despite having continued to encourage LAs to adopt RBV since 2012, why did the Department not update its ‘offer’ and guidance when the law changed?
- Did DWP make any assessment of the potential liabilities to which its policy recommendations exposed volunteer LAs?

Q2: If DWP is itself using RBV or automated risk assessment algorithms anywhere in UC processes, (a) what information does it provide to individuals about this processing; and (b) what *simple* way has it introduced to allow claimants to request human intervention or challenge a decision, as per the law and ICO guidance¹¹?

The arbitrary nature of RBV / automated risk assessment algorithms can ‘cut holes’ in the social safety net with no warning, explanation or recourse for those individuals who fall through them – leaving questions that remain entirely unanswered by the State.

RBV is being used...

While not necessarily always labelled or referred to as “Risk Based Verification”, some form of targeting is being used in the social safety net – both inside DWP for Universal Credit, and out beyond the DWP ‘boundary’ at local level, i.e. for Housing Benefit and Council Tax Reduction / Council Tax Support in local authorities.

Though an attempt has been made to ‘outsource’ RBV to local government for Housing Benefit and Council Tax Support, automated risk assessment is increasingly a part of the wider benefit system – incorporating algorithm-based decision-making tools that are affecting a wider and wider range of benefit claimants, including claimants for Universal Credit.

DWP’s 2018 annual report plainly states that “Artificial intelligence is enhancing our efforts to tackle benefit fraud”,¹² at least some of which is clearly focused on risk assessment:

We continue to trial new external datasets to risk assess the accuracy of UC claims and reduce what claimants and agents have to prove when they make a claim. A two year contract with an external credit reference agency introduced new data and is identifying potential ‘living together’ fraud.

And, with specific regard to Universal Credit and risk assessment, the National Audit Office’s 2018 report on Universal Credit¹³ indicates that work to develop a “fully automated risk analysis and intelligence system on fraud and error” continues:

¹¹ e.g. <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to-automated-decision-making-including-profiling/>

¹² page 63, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/721224/dwp-annual-report-and-accounts-2017-2018.pdf

¹³ page 61, <https://www.nao.org.uk/wp-content/uploads/2018/06/Rolling-out-Universal-Credit.pdf>

3.29 Fraud and error systems and reports are not yet complete. The Department intends to develop a fully automated risk analysis and intelligence system on fraud and error ... from April 2019 onwards.

RBV has, in practice, significantly changed how 'High Risk' applicants and claimants are treated.

Risk-Based Verification manifests policy reasoning as follows: everyone is treated as 'High Risk' by default, and then there can be an 'easement' of those conditions, i.e. depending on the profiling of a hidden algorithm, some people will be required to show fewer documents, or will be allowed to provide photocopies not originals. The policy is justified by moving 55% of people into the 'Low Risk' category.¹⁴

However, while prior to the introduction of RBV everyone was treated as 'High Risk' and had to show original documents, as a *direct result* of DWP's insistence that the use of RBV has to be "cost neutral",¹⁵ the operationalisation of RBV has materially changed how those who are now profiled as 'High Risk' are treated: a raft of new measures and processes are applied to the highest risk group.

The resources saved (staff time, etc.) by *not* checking 55% of applicants as rigorously as before are being targeted entirely at those deemed 'High Risk' – 75% of fraud resources are targeted against the 20% of applicants deemed 'High Risk', with the remaining 25% of resources are targeted at the 20% deemed 'Medium Risk'.

In practice, the "20%" deemed to be 'High Risk' has actually been as low as 13%, or potentially lower,¹⁶ resulting in *seven times* more scrutiny of those deemed 'High Risk' than other groups. And the measures targeted against 'High Risk' claimants could not be applied to everyone, simply due to the level of resourcing that would be required to do so.

As illustrated by the table on the next page, one council has broken down the risk outcomes according to various characteristics,¹⁷ with those considered the greatest risk being older people who do not receive DWP benefits, then those living in council housing or private accommodation, and then those who are of working age. (Noting that those who are on UC will have been 'passported' for approval via DWP/UC.)

¹⁴ Table on page 7: <https://democracy.west-norfolk.gov.uk/documents/s3445/Housing%20Benefit%20and%20CT%20Support%20Risk%20Based%20Verification.pdf>

¹⁵ page 3, paragraph 10: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/637068/g1-2016.pdf - "The introduction of RBV by an LA should be cost neutral, i.e. better direction of resources towards high risk cases and **not less resource being utilised.**"

¹⁶ <https://democracy.leeds.gov.uk/documents/s65921/RBV%20report%20AG%20Committee%20March%202012.pdf>

¹⁷ <https://www2.harrow.gov.uk/documents/s161655/RBV%20GARMS%20report%20Jan%202020%20v2%20-%20final.pdf>

2.20 Breakdown of caseload across risk categories

2.21 Limited information is currently available on the claim types within each risk group as claims need to be assessed for this data to be extracted. The data below is for claims received in November 2015

Table 3 - Claim types across risk categories November 2015

Claim type	% low risk	% medium risk	% high risk
Passported* working age	73%	23%	4%
In-work working age	42%	41%	17%
PCGC pensioner	100%	0%	0%
Non-passported pensioner	50%	17%	33%

*Automatically passported entitlement onto Housing Benefit/Council Tax Support due to receipt of another state benefit

Table 4 - Tenure type within risk categories November 2015

Claim type	% low risk	% medium risk	% high risk
Private	39.3%	45.3%	15.4%
Council	70.0%	10.0%	20.0%
Housing Association	88.9%	5.6%	5.5%

2.22 This caseload data demonstrates that more complex cases are most likely to require some level of verification of entitlement, and those on passported benefits or Social Rented accommodation are more likely to fall into the low risk category. The Risk Based Verification process allows resources to be directed toward those more complex cases which have a much higher probability of being correct and away from those simpler cases.

This same council later removed the use of RBV for housing benefit and Council Tax support,¹⁸ based on an Equality Impact Assessment, which stated:¹⁹

“Due to the complexity of the supported and temporary accommodation cases they are normally treated as being medium or high risk and therefore do not benefit from the Risk Based Verification process.” [i.e. their checks would not be reduced from normal]

Measures taken against these more ‘complex cases’ include Credit Reference (CRA) checks, enhanced CRA checks, other forms of cross-referencing (e.g. land registry checks), voice-recorded telephone interviews, visits to applicants’ homes, and periodic follow-ups.

Q5: Has DWP calculated the cost and impact of applying the current ‘High Risk’ measures that its “cost neutral” RBV policy has introduced, when or if they were to be (re)applied to all?

- If RBV were to be found unlawful for any reason then – as a consequence, and as per policy – every UC and benefit claimant would once more have to be classified as ‘High Risk’ by default, with no easements.

Q6: Do those local authorities currently using commercial RBV solutions, as recommended by DWP, have the staff and resources to be able to administer Housing Benefit and Council Tax Support, should they be forced to return to treating all claimants as ‘High Risk’?

¹⁸ <https://www2.harrow.gov.uk/documents/s161655/RBV%20GARMS%20report%20Jan%202020%20v2%20-%20final.pdf>

¹⁹ <https://www2.harrow.gov.uk/documents/s161577/RBV%20EQIA%20Feb%202020.pdf>

Some councils (e.g. Redditch²⁰) are apparently already ‘overriding’ the RBV algorithm, to the extent that they *always* classify claims for certain things – e.g. “exempt accommodation”, where the rules that normally limit the amount of rent covered by a Housing Benefit award or UC do not apply – as being ‘High Risk’.

Q7: To what extent have local ‘exemptions’ to RBV business rules been assessed for legality and proportionality? Does DWP require that any deviation from the pre-determined risk models be reported, or justified?

This is relevant because while such blanket escalations may be a local authority’s prerogative, given that the discretion permitted to their assessment officers allows those officers to only escalate by a *single* level of risk (i.e. from ‘Low’ to ‘Medium Risk’, or from ‘Medium’ to ‘High Risk’) and that at least some of the people making, e.g. “exempt accommodation” claims would be considered ‘Low Risk’ by the RBV algorithm – if exposed to it – deeming such people to be ‘High Risk’ (i.e. automatically bumping them up *two* levels, with no right of challenge or appeal) would seem to be open to challenge.

²⁰ Final paragraph, page 2:

<https://moderngovwebpublic.redditchbc.gov.uk/documents/s21783/Risk%20based%20varification%20policy%20report%20-%20appendix.pdf>

Background to RBV, and its interactions with the benefit system

HB/CTB Circular S11/2011

Risk-Based Verification of HB/CTB Claims Guidance

Introduction

1. This guidance outlines the Department's policy on Risk-Based Verification (RBV) of Housing Benefit and Council Tax Benefit (HB/CTB) claims.

Background

2. RBV allows more intense verification activity to be focussed on claims more prone to fraud and error. It is practiced on aspects of claims in Jobcentre Plus (JCP) and the Pension Disability and Carers Service (PDCS). Local authorities (LAs) have long argued that they should operate a similar system. It is the intention that RBV will be applied to all Universal Credit claims.
3. Given that RBV is practised in JCP and PDCS, the majority (up to 80%) of HB/CTB claims received in an LA may have been subject to some form of RBV. Already 16 LAs operate RBV. Results from these LAs have been impressive. In each case the % of fraud and error identified has increased against local baselines taken from cells 222 and 231 of the Single Housing Benefit Extract (SHBE). In addition, in common with the experience of JCP and PDCS there have been efficiencies in areas such as postage and storage and processing times have improved.
4. We therefore wish to extend RBV on a **voluntary basis** to all LAs from April 2012.

The Department for Work and Pensions' 2011 Housing Benefit and Council Tax Benefit Circular²¹ defines risk-based verification as follows:

RBV is a method of applying different levels of checks to benefit claims according to the risk associated with those claims. LAs will still be required to comply with relevant legislation (Social Security Administration Act 1992, section 1 relating to production of National Insurance numbers to provide evidence of identity) while making maximum use of intelligence to target more extensive verification activity on those claims shown to be at greater risk of fraud or error.

As per the extract on the previous page, the explicit intent of RBV is thus for "**more intense verification activity to be focussed on claims [deemed to be] more prone to fraud and error**".

The adoption of RBV by local authorities to profile applicants applying for Housing Benefit (HB) and Council Tax Support (CTS) began on a voluntary basis in April 2012. Local authorities were required to establish a "*robust baseline against which to record the impact of RBV*", though the source of this baseline was left to the LA to determine.

21

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/633018/s11-2011.pdf

Notably, though DWP stated itself to be “*of the opinion that the use of IT will support the success of RBV, it does not in any way endorse any product or company*”, this places local authorities that volunteer to implement RBV in an invidious position, because:

Q8. If DWP neither endorses any product nor company, nor provides local authorities with copies of the models it developed in its two year ‘research period’ (prior to 2011) against which any product can be tested; and if the commercial providers of RBV software will not reveal their algorithms to the LAs, then how can LAs – as data controllers – in any meaningful way be determining “the purpose and means of personal data processing” as per DPA 2018 and GDPR?

In contrast to, say, NHS ‘GP Systems of Choice’ (GPSoC)²² – the contractual framework to supply IT systems and services to GP practices across England, and to support practices in procuring what they need to meet their legal and other obligations – DWP’s approach is so ‘hands-off’, allusive, unsupportive, unaccountable, and unclear as to seem borderline underhand.

Encouraging local authorities to “volunteer” themselves, and the citizens whom they serve, to use (or be ‘used’ by) systems that lack even basic transparency, while handing off responsibility for implementing *inherently discriminatory* automated processes to a market of IT suppliers and credit reference agencies, with no oversight or audit, and only lip service to legal considerations opens both DWP and those LAs who do adopt RBV to significant risk.

Q9. If the processing and business rules of these commercial ‘black box’ RBV algorithms can only be inferred by looking at their outputs (i.e. from how many people are put into each risk category) and guidance from DWP already specifies that “no more than 55% of cases should be Low Risk, 25% Medium Risk and 20% High Risk”²³ then how meaningful are the actual checks?

- Given that RBV algorithms can be (and are) tuned to deliver the desired targets, can what is being ‘measured’ even be considered ‘risk’? It would seem that *any* algorithm delivering the expected proportions could be deemed to be operating ‘correctly’...

Q10. If local authorities cannot determine the business rules being applied, or adjust how those rules are applied to different groups, how are they in any position to meet their obligations under the Equality Act 2010?

- If an LA cannot determine how, or even if, the RBV algorithms are taking account of Protected Characteristics, or decide which characteristics are being processed and what weightings (if any) are being applied to them, then how – other than very carefully designed *post hoc* sampling – can any LA determine whether people with Protected Characteristics are *not* being discriminated against?

²² <https://digital.nhs.uk/services/gp-systems-of-choice>

²³ e.g. paragraph 2.6,

<https://rbwm.moderngov.co.uk/documents/s25679/RBV%20report%20for%20APRP.pdf>

Credit reference agencies

It appears that, in practice, while DWP does not *formally* endorse any particular company's RBV product or service, most LAs tend to use one of a small number of suppliers – of which several cloud-based providers appear on the GOV.UK Digital Marketplace.²⁴

Building on FOI work by Big Brother Watch,²⁵ The Bureau of Investigative Journalism (TBIJ), the Guardian and others, it is clear that the larger commercial IT providers such as Capita, Civica and Northgate provide RBV software to local authorities through modules in, or integrated with, their existing local government platforms.

One of the more significant players in the RBV market is a company called Coactiva – which, having partnered with the credit reference agency Equifax in 2012, was bought by another credit reference agency, CallCredit (now TransUnion), in 2014. Coactiva's RBV software appears under various brands of TransUnion's 'ThreeSixty RBV' offering on the GOV.UK Digital Marketplace.²⁶

Another supplier known to be used by multiple local authorities is a company called Xantura,²⁷ which sells a risk model it claims "*utilises around 50 variables to predict the likelihood of fraud and error*" – and which went so far as to ask the question on its own corporate blog on 14 May 2019, "*Is RBV still relevant after UC roll out?*"²⁸

A major investigation by TBIJ in 2019²⁹ showed that, while IT contracting for government is "*still dominated by traditional big names, many lesser known companies are also offering a wide variety of services*". TBIJ discovered over 1,800 companies providing such services, though

"many authorities were unwilling or unable to specify how and why they purchased these services, however, or what their precise specifications were."

Given this lack of clarity – despite DWP's 'doubling down'³⁰ – and in light of growing numbers of local authorities that are dropping RBV providers and, in some cases, RBV altogether:

²⁴ e.g. a search for "Risk Based Verification Council Tax" returns five results:

<https://www.digitalmarketplace.service.gov.uk/g-cloud/search?q=Risk%20Based%20Verification%20Council%20Tax&lot=cloud-software>

²⁵ <https://www.ohchr.org/Documents/Issues/EPoverty/UnitedKingdom/2018/NGOS/BigBrotherWatchAnnex1.pdf>

²⁶ <https://www.digitalmarketplace.service.gov.uk/g-cloud/services/730776040571418> and <https://www.transunion.co.uk/industry/government>

²⁷ <https://xantura.com/what-we-do/debt-and-benefits/>

²⁸ <https://web.archive.org/web/20200519160449/https://xantura.com/is-rbv-still-relevant-after-uc-roll-out/>

²⁹ <https://assets.documentcloud.org/documents/5993565/2019-05-08-TBIJ-Government-Data-Systems-Published.pdf>

³⁰ DWP's 2016 "cost neutral" RBV reminder circular: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/637068/g1-2016.pdf

Q11. Does DWP maintain a (central) list of how many local authorities have implemented RBV, using which software, at what cost and to what end result?

- **Are LAs required to report to DWP that they are ‘volunteering’ to use RBV? If so, what do they have to report? If not, why not?**

Q12. On the evidence, DWP’s policy of pushing RBV ‘out to the edges’ of UC, relying solely upon market provision, with minimal guidance or support for LAs, has failed to deliver value in more than one area. What processes did DWP put in place to ensure that its ‘offer’ to volunteer LAs was going to deliver, (a) the stated policy intent, and (b) public value?

- **If this information is not (readily) available, it would appear that DWP has been pursuing an automation agenda *without even wanting to know the outcomes*. Who will hold DWP accountable for its policy failure?**

RBV is also indicative of government's approach to fraud, i.e. it appears to be more about political appearances and risk-shifting, rather than anything that makes a substantive contribution (or saving) to the public purse – and which, while making some things marginally easier for around half of applicants, puts significant increased burdens onto the rest.

Q13. Given the introduction of RBV across UC was DWP’s stated intent in 2011, how far advanced are its plans to incorporate automated risk assessment across its central / regional / local ‘Fraud & Error’ functions, and/or anywhere else in the system?

- **How much taxpayers’ money has DWP spent on developing and assessing RBV and related technologies, including paying for access to data held by credit reference agencies?**

What does the RBV process look and feel like to a claimant?

The claimant is only able to infer the result of the RBV process from what they are asked to provide to the assessor; must they provide original documents, copies, or nothing at all?

If someone is put into the 'Low risk' category then checks should be minimal. Similarly, for those assigned to the 'Medium risk category', there will generally be a degree of 'leeway' in the evidence they are asked to produce; for example, photocopies or scans would be considered acceptable, rather than original documents

Those the RBV software puts into the 'High Risk' category will feel the brunt of the suspicion and bureaucracy. They will have to provide original documents, not copies, for most things and will be subject to other measures that – while these vary from LA to LA – may include:

- checks with credit reference agencies / enhanced CRA checks;
- checks via the 'National Anti-Fraud Network' (NAFN);
- other forms of cross-referencing, e.g. land registry checks;
- voice-recorded telephone interviews;
- in-person interviews and visits to applicants' homes;
- periodic follow-ups by phone, text or in person.

For UC itself, if DWP's own risk-based verification deems a claimant to be 'High Risk', while that person may successfully pass GOV.UK Verify's identity assurance process, DWP will still require them to present original documents (e.g. at a JobCentre) – leading them to assume that they failed the Verify process [[see Annex 2A](#)].

What have we learned that can help people?

DWP requires³¹ the audit committee of each local authority to approve its Risk Based Verification Policy and practices, although some of the information – relating to ‘risk categories’ – may be redacted.

Even though their risk category is not disclosed to a claimant, the information shown in the tables on the following pages would allow claimants, in most cases, to determine the risk category to which their local authority’s RBV software has assigned them, simply through the types of documents they are asked to provide.

Evidence Type	Subcategory	Low Risk	Medium Risk	High Risk
Identity & NINO	Claimant’s ID/NINO	> Originals. CIS Check where identity has been verified for a qualifying benefit	> Originals, CIS Check where identity has been verified for a qualifying benefit	> Originals Required, CIS Check
Residency & Rent	> Private Tenants	>	> Originals or Photocopies;	> Originals Required
	> Social Landlords	>	> Originals, Photocopies or electronic file from Social Landlord	> Originals Required or electronic file from Social Landlord
	> Registered	>	> Originals, Photocopies	> Originals Required
Household	> Partner’s ID/ NINO	> Originals or CIS check where identity has been verified for a qualifying benefit	> Originals or Photocopies CIS check	> Originals Required or CIS check
	> Dependents (responsibility for, not ID of dependants)	>	> Originals or Photocopies CIS check	> Originals Required or CIS check
	> Non-Dependent Working	>	> Originals or Photocopies or WuRTI, RTI file or VEP alert	> Originals Required or WuRTI, RTI file or VEP alert
	> Non-Dependent (P8) Non Dependent no Income	>	> CIS Check P45 or statement	> CIS Check P45 or statement
	> Non-Dependent Student	>	> Originals or Photocopies	> Originals Required
	> 2AR: Non-Dependents Not Working	>	> Originals or Photocopies	> Originals Required
Income	> State Benefits	>	> CIS Check or Originals or Photocopies	> CIS Check or Originals or Photocopies
	> Earnings, SSP, SMP & SPP	>	> Originals, Photocopies or Cert of Earnings or WuRTI, RTI file or VEP alert	> Originals Required or WuRTI, RTI file or VEP alert
	> Self Employed	> SE Pro-Forma	> Originals, Photocopies, SE form or audited accounts	> Originals Required – Receipts and Invoices
Child Care Costs		>	> Originals or Photocopies	> Originals Required
Students	> (Income + Status Required)	> Photocopies	> Originals or Photocopies	> Originals Required
Capital	> Working Age Working Age & > £6,000	>	> Originals or Photocopies	> Originals Required; must include last 2 months’ transactions
	> Elderly Elderly & > £10,000	>	> Originals or Photocopies	> Originals Required; must include last 2 months’ transactions
	> Property	> LAREV1	> Originals, Photocopies and LAREV1	> Originals Required and LAREV1

Cases in High Risk Group will also be subject to a credit check via NAFN within one month of the date of the claim unless the claim has ended prior to the check being conducted.

When going through an official process, the ‘user’ (citizen, claimant...) will be given some information. Whether that is on screen, on paper forms, on leaflets and support materials, or in related guidance, this information is rarely enough to explain the system in and of itself – *but it is the only thing that a citizen can be expected to have seen.*

Therefore, when different people are told – or asked for – different things by officials, there is likely a reason behind that.

³¹ As required by DWP in paragraph 4 of its 2011 Circular: “The information held in the Policy, **which would include the risk categories**, should not be made public due to the sensitivity of its contents.”

Appendix A

Type of Evidence	Sub-category of evidence	Low Risk	Medium Risk	High Risk	
Identity and S19	Identity	Originals or Photocopies	Originals required	Originals Required	
	S19	Originals or Photocopies accepted	Originals required	Originals Required	
Residency/Rent	Private Tenants		Originals or Photocopies accepted	Originals Required	
	Social Landlords			Originals Required	
	Public Sector				
Household Composition	Partner ID/S19/Income/Capital	Originals or Photocopies accepted	Originals or Photocopies accepted	Originals Required	
	Dependants under 18		Originals or Photocopies accepted	Originals Required	
	Non Dependants – remunerative work			Originals or Photocopies accepted	Originals Required
	Non Dependants – passported benefit				
	Non Dependant –student		Originals or Photocopies accepted	Originals Required	
	Non Dependant – not in remunerative work/other			Originals or Photocopies accepted	Originals Required
Income	State Benefits				
	Earnings/SMP/SSP	Originals or Photocopies accepted	Originals or Photocopies accepted	Originals Required	
	Self Employed Earnings	Originals or Photocopies accepted	Originals or Photocopies accepted	Originals Required	
Child Care Costs			Originals or Photocopies accepted	Originals Required	
Student Status	Income also required	Originals or Photocopies accepted	Originals or Photocopies accepted	Originals Required	
Capital	Below lower capital limit				
	Above lower capital limit		Originals or Photocopies	Originals Required	
	Property		Originals or Photocopies accepted	Originals Required	

While some such ‘reasons’ may amount to little more than arbitrary official randomness, some will be significant systemic choices; knowing *which types of documents* a claimant is required to bring as proof of Residency / Rent – or in other categories – can allow one to determine the risk category to which the RBV systems has assigned that individual with a high degree of certainty, because that is the RBV process explicitly approved by a council audit committee.

As shown in the tables above and the one on the following page, from Rochdale Borough Council,³² if a claimant is *required* to provide original documents then **they can be sure they have been assigned to the ‘High Risk’ category.**

Similarly, if they are told they need provide no documentary evidence of, e.g. their earnings, their savings or their child care costs, then they can know they have been put into the ‘Low Risk’ category. And if they are given the *option* to provide photocopies to evidence these facts, they know they have been designated as ‘Medium Risk’.

³² <http://democracy.rochdale.gov.uk/documents/s38488/Append.%201%20for%20Risk%20Based%20Verification%20Policy.pdf> - copy downloaded on 21 May 2020.

Appendix 1				
Type of Evidence	Sub-category of evidence	Low Risk	Medium Risk	High Risk
Identity and National Insurance Number (NINO) for customer and partner	Identity	Originals, Photocopies, or scanned documents	Originals, Photocopies, or scanned documents	Originals required
	NINO	Originals, Photocopies, scanned documents or CIS check	Originals, Photocopies, scanned documents or CIS check	Originals, or CIS check required
Residency/Rent	Private Tenants	No evidence required	Originals, Photocopies, or scanned documents	Originals required
	Social Landlords	Rent breakdown confirmed via electronic file held by Revenues and Benefits or obtained directly from Social Landlord	Residency confirmed by Social landlord or contact with tenant Rent breakdown confirmed via electronic file held by Revenues and Benefits or obtained directly from Social Landlord.	Residency confirmed by Social landlord or contact with tenant Rent breakdown confirmed via electronic file held by Revenues and Benefits or obtained directly from Social Landlord.
Income customer and partner	Passported Benefits	CIS check	CIS check	CIS check required
	Non-passported Benefits	CIS check	Originals, Photocopies, scanned documents or CIS check	Originals, or CIS check required
	Earnings/SMP/SSP	No evidence required	Originals, Photocopies, or scanned documents	Originals required
	Self-employed earnings	No evidence required	Originals, Photocopies, or scanned documents	Originals required
	Other	No evidence required	Originals, Photocopies, or scanned documents	Originals required
Capital customer and partner	Capital below £5,500 (working age), £9,500 (pension age)	No evidence required	Originals, Photocopies, or scanned documents	Originals required
	Capital above £5,500 (working age), £9,500 (pension age)	No evidence required	Originals, Photocopies, or scanned documents	Originals required
	Property	No evidence required	Originals, Photocopies, or scanned documents	Originals required
Household Composition	Dependants	No evidence required	Originals, Photocopies, or scanned documents	Originals required
	Non-dependants – passported benefit	No evidence required	Originals, Photocopies, scanned documents or CIS check	Originals or CIS check required
	Non-dependant - other	No evidence required	Originals, Photocopies, or scanned documents	Originals required
Child Care Costs		No evidence required	Originals, Photocopies, or scanned documents	Originals required

In addition to the forms of paperwork they are required to provide, if a claimant receives, e.g. a home visit within 6 months of making a claim – or if they discover, or are told, that they have been subject to a credit check – they can be pretty certain that they are being treated as ‘High Risk’.

This being the case, and assuming that people are now able to determine the category into which the RBV software has put them, what else does the legal community need to challenge the ‘risk scores’?

Process questions may arise as to what is legitimate to require from claimants and what is not – and what is discriminatory, but those questions can now, in a digital system, be framed in such a way as to be *measurable*.³³

For those supporting claimants, it is important to keep separate what it is that *users believe they are providing and why*, from one's own *professional understanding*, and also *what official statements say* about the system. Lack of transparency, lack of clarity or misleading information may in themselves represent actionable offences, e.g. failure to perform 'fair processing' (under GDPR, the 'right to be informed') or breaching the Data Minimisation Principle.

So, for example, DWP requires that Risk Based Verification (RBV) of, e.g. Housing Benefit claims have consistent audit processes, which are then subject to annual certification.³⁴

If those audit processes sample 10 'High Risk' claimants and they are all women (or all men, or all share a protected characteristic) that does not necessarily imply an equality breach; even if all 10 people in the high risk category were exclusively women, this would be a single *unbalanced* sample, but not an invalid one.

However, if three *independent* audit processes were each to sample 10 'High Risk' claimants, and in total those samples did not average back to the expected population, then that would provide a sufficient evidence base on which to ask a question: is the audit process flawed, or is it the system?

N.B. Even were such biases to be exposed, this does not necessarily mean that an individual LA is directly at fault – the fault most likely lies with their common supplier – neither is it necessarily a data breach. But if the DPIA of a particular LA's RBV implementation states that the Equality Act is being followed (which it should) and it turns out that the Equality Act is being breached, despite what an Equality Impact Assessment says, then regulatory concerns can be raised...

In general, when systems give out different information based on different inputs, they can *always* be detected, and *usually* tested. Every system will naturally reflect the incentives of its creator(s); this is unavoidable.³⁵

In the simplest terms, if a system does not give differential output, then it is not much use for assisting in decision-making! For algorithms, models and Machine Learning systems to be usable by a public body, they have to produce something which can be acted upon – and those actions can be constrained by law.

³³ If civil society can create screenshots for documentation, then that same process measures the outcomes of risk based verification choices, and a statistical analysis can inform lawfulness debates.

³⁴ paragraphs 18 & 19, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/633018/s11-2011.pdf

³⁵ <https://medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53>

Is RBV making problems worse?

Multiple local authorities' Equality Impact Assessments (or equivalent) claim a "neutral impact"³⁶ across age and all of the protected characteristics under the Equality Act 2010. Though some LAs do acknowledge such impacts are possible,³⁷ none we have seen arrive at the conclusion that people who share protected characteristics are unduly affected. Which seems unlikely at best...

Some councils do after-the-fact 'blind sampling' tests to check if the algorithm is performing 'to expectation', i.e. asking little more of the RBV software than "is it tending to put people into the right risk category?" – but none appear to be monitoring equality impacts in the systemic fashion identified by Dr Byrom's work (or similar). Nor do the commercial providers appear to offer the ability to do so.

None of the councils can see either DWP's models or the commercial algorithms, or how they operate; they essentially have to take DWP's and their supplier's word that the black box they bought is performing as intended. The general assertion seems to be, "Because we're doing this to everyone, no-one can possibly be being discriminated against" – a logical fallacy that is as dangerous as it is discriminatory.

If *any* protected characteristics are being fed into the algorithm,³⁸ then clearly they *could* be having an effect on the output. If the council (and DWP) do not know what those effects and/or business rules are, and how any such rules meet both the policy intention and LA's lawful obligations, then how can any substantive claim be made with regard to LAs having met their public sector duties under the Equality Act 2010?

Q14: What data fields are being fed to any / each of the RBV algorithms? Are any of them protected characteristics?

³⁶ e.g. Ealing Council's 2018 RBV Policy disregards possible impacts on those with Protected Characteristics, on the basis that any impact should be "neutral" *simply because RBV is applied to all applicants*: <https://ealing.cmis.uk.com/ealing/Document.ashx?czJKcaeAi5tUFL1DTL2UE4zNRBcoShgo=195YowuP8la3ZeGBVU2bAkxQ2iVpofBRpzm%2FXRC%2BMEigYSxJ52Gbmw%3D%3D&rUzWRP%2BZ3zd4E7lkn8Lyw%3D%3D=pwRE6AGJFLDNih225F5QMaQWCtPHwdhUfCZ%2FLUQzgA2uL5jNRG4jdQ%3D%3D&mCTIbCubSFfXsDGW9IXnlq%3D%3D=hFflUdN3100%3D&kCx1AnS9%2FpWZQ40DXFvdEw%3D%3D=hFflUdN3100%3D&uJovDxwdjMPoYv%2BAJvYtyA%3D%3D=ctNJFf55vVA%3D&FgPIIEJYlotS%2BYGoBi5olA%3D%3D=NHdURQburHA%3D&d9Qjj0ag1Pd993jsyOJqFvmyB7X0CSQK=ctNJFf55vVA%3D&WGewmoAfeNR9xqBux0r1Q8Za60lavYmz=ctNJFf55vVA%3D&WGewmoAfeNQ16B2MHuCpMRKZMwaG1PaO=ctNJFf55vVA%3D>

³⁷ e.g. Harrow Council's 2013 Equality Impact Assessment, while accepting that people with Protected Characteristics may be affected by RBV – and despite stating "*It is not possible, in advance of the policy being implemented, to identify the groups of people that fall within low, medium or high risk categories*" – makes the blanket assertion that its implementation of RBV will have no "*disproportionate adverse impact on any of the Protected Characteristics*". <https://www2.harrow.gov.uk/documents/s128005/Appendix%20B%20EqIA.pdf>

³⁸ See, e.g. Medway Council's 2017 Diversity Impact Assessment, which claims "*The [RBV] model does not take into account any of the protected characteristics dealt with by the Equalities Act*". Is this true? How do they know? And how can they determine that those with Protected Characteristics are *not* being adversely affected by an algorithm that, while it may not be directly fed such characteristics – are they really claiming no information about age is entered, and no sex? – is likely able to operate on proxies for those characteristics: <https://democracy.medway.gov.uk/mgconvert2pdf.aspx?id=36387>

- **If yes, then how are they ensuring that the algorithms are meeting statutory requirements and public sector duty?**
- **If no, then how can the councils / DWP know what the impact is on these groups? Just because you choose not to / don't measure something doesn't mean that it's not going on...**

Legal opinion commissioned by The Legal Education Foundation (TLEF) from Robin Allen QC and Dee Masters in 2019³⁹ suggests that:

*Whilst local authorities are undertaking Equality Impact Assessments, the ones which we have seen have tended to be very superficial. It is not fanciful to imagine that the **RBV processes which are being deployed by local authorities might be acting in a discriminatory way**. After all, there is some publicly available data which demonstrates that RBV schemes can act in surprising ways, for example, identifying high numbers of women as being at higher risk of committing fraud. Equally, the House of Commons Science and Technology Select Committee noted, as early as 2018, how **machine learning algorithms can replicate discrimination**.*

Regardless of the particular models or technologies being used to automate the decision-making process, the routine use of RBV software may introduce and embed what is known as “automation bias”⁴⁰ – where ‘wrong information’ in a system tends to be presumed correct, and where officials become ‘complacent’ about a system that appears to be working as expected.

Testing RBV against ‘value for money’

Bearing in mind the answer(s) to Q9, several local authorities that volunteered to use RBV have – after several years of using it – done their own assessments of the approach, and are arriving at similar conclusions

In its annual audit, Manchester City Council performed an analysis of the impact of RBV within its own processes and decided, in March 2019,⁴¹ to cease using RBV in its administration of Housing Benefit and Council Tax Support. It did so despite the encouragement of DWP and funding ‘incentives’ from the Ministry of Housing, Communities and Local Government (MHCLG), e.g. on budgetary efficiency, and the requirements of the Prevention of Social Housing Fraud Act 2013.

³⁹ In particular, pages 40-54: <https://www.cloisters.com/wp-content/uploads/2019/10/Open-opinion-pdf-version-1.pdf>

⁴⁰ There is a significant body of research on automation bias in various domains, going back decades, e.g. ‘Accountability and automation bias’ in the International Journal of Human-Computer Studies, April 2000: <https://skitka.people.uic.edu/JHCS2000.pdf> and more recently, e.g. this systematic review from 2012: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3240751/>

⁴¹ Screengrab is taken from <https://democracy.manchester.gov.uk/ieListDocuments.aspx?CId=144&MID=341>

And, in this instance, the City Council saved £14,000 per annum:

16. Risk Based Verification

- [View the reasons why item 16. is restricted](#)

A report of the Director of Customer Services and Transactions is enclosed.

Minutes:

Members considered the report the City Treasurer proposing changes to the verification of Housing Benefit and Council Tax claims.

Decision

1. To agree that the Council cease to use Risk Based Verification in the administration of Housing Benefit and Council Tax Support from 1 April 2019.
2. To agree that the Council revert use verification standards, not aligned to Risk Based Verification, which meet Department of Work and Pensions requirements for Housing Benefit cases and associated Council Tax Support Cases.
3. To agree that for other than Council Tax Support claims, the Council should, where appropriate, apply a lighter touch approach to verification while taking judicious steps to prevent fraud and error that reduce the risk of excessive remission of Council Tax charges.
4. To agree that the Council cancel the contract with the external supplier that supports Risk Based Verification at a saving of £14000 per annum.

Another positive is that this does suggest ‘outsourcing’ automation to councils (with incentives, positive or negative) can provide at least some evidence-based, independent and democratic involved oversight of whether such approaches work.

Following Manchester in March, having first introduced it in 2015, in September 2019 North Tyneside Council stopped using the RBV system it had bought from Coactiva (now Transunion),⁴² citing both a reduction in the number of claims it was processing and a “significant improvements in the access to data provided by the Department for Work and Pensions”. In January 2020, Harrow Council also proposed “to remove Risk Based Verification from the administration of Housing Benefit and Council Tax Support with effect from 1st April 2020”,⁴³ citing similar reasons.

Q15. What would happen were, e.g. one of the political parties to ask every local authority using RBV to produce the assessment that DWP won’t do itself?

With regard to overblown promises of automation in the public services more generally, an investigation by the Guardian newspaper in 2019⁴⁴ suggests that Hackney Council – while not specifically mentioning RBV – was dissatisfied with Xentura’s predictive modelling, which “wasn’t able to provide sufficiently useful insights”. And at least part of the reason why Sunderland City Council chose to end the £4.5 million “Intelligence Hub” it had commissioned Palantir to build in 2014 was that it could not deliver the required outcomes.

⁴² <https://my.northtyneside.gov.uk/sites/default/files/meeting/related-documents/5c%20Risk%20Based%20Verification%20Report.pdf>

⁴³ <https://www2.harrow.gov.uk/documents/s161655/RBV%20GARMS%20report%20Jan%202020%20v2%20-%20final.pdf>

⁴⁴ <https://www.theguardian.com/society/2019/oct/15/councils-using-algorithms-make-welfare-decisions-benefits>

For those who end up ‘picking up the pieces’ of RBV failures, Dr Byrom’s 13 data items, as tabulated on the following page,⁴⁵ provide a measure to detect bias in processes:

Figure 0-2 **Minimum data required to measure vulnerability**

Individual attributes to be captured	Data in practice directions used to identify vulnerability?	Data related to digital exclusion?	Data needed to monitor duties under the Equality Act 2010?
1 Age	✓	✓	✓
2 Disability ¹	✓	✓	✓
3 Employment status/income	✓	✓	✗
4 English as a foreign language	✓	✗	✗
5 Gender reassignment	✗	✗	✓
6 Highest level of education (proxy for literacy)	✓	✓	✗
7 Postcode (permanent address, to identify whether in a care home, homeless, in an area of low internet coverage etc.)	✓	✓	✗
8 Pregnancy and maternity	✗	✗	✓
9 Race	✗	✗	✓
10 Religion or belief	✗	✗	✓
11 Sex	✗	✗	✓
12 Sexual orientation	✗	✗	✓
13 Fear or distress connected with the case e.g. domestic violence/abuse, in detention, survivor of trafficking/trauma	✓	✗	✗

And the (ongoing) JCWI ‘Right to Rent’ case⁴⁶ suggests ways in which the necessary judgements can be won at Judicial Review.

On realising the Home Office was not properly monitoring or evaluating its ‘Right to Rent’ policy,⁴⁷ the Joint Council for the Welfare of Immigrants chose to independently measure its effects using online and offline surveys.⁴⁸ The analysis of these results, published in JCWI’s “No Passport Equals No Home” report in 2015,⁴⁹ led to an initial ruling against the Home Office that the policy was discriminatory – a finding that was upheld at Appeal in early 2020,⁵⁰ though the judge chose to leave it to MPs and Parliament to decide whether the racial discrimination caused is “greater than envisaged”.

Algorithmic assumptions about ‘fraud’ and ‘risk’

Is it really a *risk* assessment if the algorithm is tweaked to deliver target percentages in each ‘group’? Pure risk assessment would likely be on a bell curve; shifting this implies a degree of targeting, i.e. (potential) discrimination.

⁴⁵ page 5, <https://www.thelegaleducationfoundation.org/articles/the-legal-education-foundation-is-today-publishing-a-blueprint-for-digital-justice>

⁴⁶ <https://www.bailii.org/ew/cases/EWHC/Admin/2019/452.html>

⁴⁷ <https://www.freemovement.org.uk/inspector-right-to-right-scheme/>

⁴⁸ JCWI surveyed both landlords and agents: <https://www.surveymonkey.co.uk/r/GZPNWVN> and tenants and lodgers: <https://www.surveymonkey.co.uk/r/HLF7Y5B>

⁴⁹ <https://www.jcwi.org.uk/Handlers/Download.ashx?IDMF=5ab15f33-ae32-4e66-b040-402b71392177>

⁵⁰ <https://www.jcwi.org.uk/news/court-of-appeal-agrees-that-the-right-to-rent-scheme-causes-racial-discrimination>

What are the actual models and algorithms being used? RBV software is often referred to as a “propensity” model, and statements such as:

*The model **does not take into account any of the protected characteristics dealt with by the Equalities Act but rather looks at previous claim history and credit agency information to determine the likelihood of it being fraudulent or erroneous.***

It seems clear that DWP and some local authorities are hiding behind commercial confidentiality to conceal the precise definition, the business rules and the operation of RBV algorithms – and, in so doing, dodging accountability.

Q 13. Given that “credit agency” checks are not done until after someone has been deemed ‘High Risk’, precisely what checks are being done, what fields of data are being fed to the algorithm(s), and from what source(s) to designate someone as high risk? Are these consistent across all of the RBV providers, at least one of which is itself a credit reference agency?

Is RBV necessary and proportionate?

The Hague District Court’s judgement⁵¹ in February 2020 on *Systeem Risico Indicatie* (SyRI),⁵² determined that the legislation enabling that risk profiling system does not comply with Article 8 of the European Convention on Human Rights (ECHR). Given the close functional similarities between SyRI in particular and RBV software as it is being applied in and around UC (and more broadly) in the UK, it is legitimate to ask whether the use of RBV does in fact violate British law.

While other will be more competent to pursue this legally, the essence of the argument is this: automated risk checks can only meet the conditions of Article 8 of the Human Rights Act 1998⁵³ if they are in accordance with the law, which includes ‘proportionate’ in a democratic society – for example, any and all data-based checks must satisfy the requirements of the Data Protection Act 2018. Therefore, in line with the First Data Protection Principle,⁵⁴ if a system is not lawful, fair and transparent, it has also failed to respect basic human rights. (The same, of course, applies for Article 14 HRA and the Equality Act 2010.)

The Information Commissioner’s Office published draft guidance in December 2019,⁵⁵ that includes what “lawful, fair and transparent” means with regard to explaining decisions made with AI (which is broadly defined) under GDPR, as incorporated into DPA 2018; helpfully unpacking what must be communicated to data subjects on the use and operation of algorithms.

⁵¹ English translation of full judgement:

<https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878>

⁵² SyRI is a legal instrument used by the Netherlands government to prevent and combat fraud in relation to social security and income-dependent schemes, and also taxes, social security and labour laws. SyRI is a technical infrastructure with associated procedures by which personal data can be linked and analysed anonymously in a secure environment, so that ‘risk reports’ can be generated.

⁵³ <http://www.legislation.gov.uk/ukpga/1998/42/schedule/1/part/I/chapter/7>

⁵⁴ <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/lawfulness-fairness-and-transparency/>

⁵⁵ <https://ico.org.uk/media/about-the-ico/consultations/2616434/explaining-ai-decisions-part-1.pdf>

In these cases, the GDPR requires that you:

- are proactive in "...[giving individuals] meaningful information about the logic involved, as well as the significance and envisaged consequences..." (Articles 13 and 14);
- "... [give individuals] at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision." (Article 22); and
- "... [give individuals] the right to obtain... meaningful information about the logic involved, as well as the significance and envisaged consequences..." (Article 15) "...[including] an explanation of the decision reached after such assessment..." (Recital 71)

The GDPR's recitals are not legally binding, but they do clarify the meaning and intention of its articles. So, the reference to an explanation of an automated decision after it has been made in Recital 71 makes clear that such a right is implicit in Articles 15 and 22. You need to be able to give an individual an explanation of a fully automated decision to enable their rights to obtain meaningful information, express their point of view and contest the decision.

But even where an AI-assisted decision is not part of a solely automated process (because there is meaningful human involvement), if personal data is used, it is still subject to all the GDPR's principles. The GDPR principles of fairness, transparency and accountability are of particular relevance.

With particular regard to fairness and transparency, the guidance states:

*"If an AI-assisted decision is made about someone **without some form of explanation of (or information about) the decision**, this may limit their autonomy and scope for self-determination. **This is unlikely to be fair**"*

and:

*"It is **unlikely to be considered transparent** if you are not open with people about **how and why an AI-assisted decision about them was made, or where their personal data was used to train and test an AI system.**"*

On these bases alone, in the way RBV has been implemented by many local authorities – according to DWP's explicit recommendations, such as concealing the risk categories – would appear to be open to challenge. While there may be a proper legal basis to use RBV, this does not obviate the need to comply with all other relevant laws when processing what is clearly personal data.

Quite aside from bringing into question the legitimacy of RBV as it is currently being used by government in the UK, in failing to meet all of their legal obligations – not least being fair and transparent – individual organisations and authorities risk formal action that includes mandatory audits, orders to cease processing of personal data, and potentially significant fines.

These are but two examples and many others, including the UN Rapporteur on Extreme Poverty, who submitted an *amicus curiae* brief in the SyRI case,⁵⁶ have underlined the many dangers and “severe human rights problems that emerge when welfare states turn into ‘*digital welfare states*’.”

Intellectual Property, Algorithms, and Artificial Intelligence

Local authorities were not given access to DWP’s models, nor to the algorithms in the commercial RBV products that DWP encouraged them to use, but – in order for DWP to be able to ‘approve’ or ‘recommend’ their use – these commercial products must have at least been able to reliably *match the outputs*, if not actually incorporate, the models used by DWP during its 2009-11 development period:

Q1: Does DWP (still) own its own RBV models and/or algorithms?

- **If it does, who holds the Intellectual Property Rights (IPR) to them?**
- **Is the IPR shared with any other organisations, e.g. credit reference agencies?**

Q2: Given DWP clearly developed “risk models” prior to 2011, has it incorporated any of these models / algorithms / code into any (other) Universal Credit systems, as was its stated intention in 2011?

- **Is RBV still being used in Jobcentres, e.g. for UC claimant in-person checks, or has the risk assessment function been replaced by (DWP’s) newer developments?**
- ***Non-UC supplemental: does the successor to the Pension, Disability and Carers Service, i.e. The Pensions Advisory Service (TPAS) still use RBV?***

While DWP continues to conceal its own models – not only from the public,⁵⁷ but from the very local authorities it is encouraging to use RBV – and while commercial RBV providers hide their algorithms behind the veil of ‘commercial confidentiality’, it is difficult to see how the use of such systems can be proportional, fair or even lawful.

Beyond the current generation of RBV software, the Department is itself already “testing artificial intelligence to judge the likelihood that citizens’ claims about their childcare and housing costs are true when they apply for benefits”⁵⁸ and admits to “developing this solution with our in-house Digital and Analytics teams using freely available open source machine learning software.”

⁵⁶ <https://www.ohchr.org/Documents/Issues/Poverty/Amicusfinalversionsigned.pdf>

⁵⁷ This FOI response in January 2019 is typical of DWP when asked about its development and use of algorithms / AI technologies: <https://www.whatdotheyknow.com/request/533365/response/1291183/attach/html/2/FOI2018%2010171%20redraft%20CFIT.pdf.html>

⁵⁸ <https://www.theguardian.com/technology/2019/oct/14/fears-rise-in-benefits-system-automation-could-plunge-claimants-deeper-into-poverty>

DWP's 2018 annual report plainly states that "Artificial intelligence is enhancing our efforts to tackle benefit fraud",⁵⁹ at least some of which is clearly focused on risk assessment:

We continue to trial new external datasets to risk assess the accuracy of UC claims and reduce what claimants and agents have to prove when they make a claim. A two year contract with an external credit reference agency introduced new data and is identifying potential 'living together' fraud. Work with other data providers is exploring how we can identify undisclosed capital.

And, with specific regard to Universal Credit and risk assessment, the National Audit Office's 2018 report on Universal Credit⁶⁰ indicates that work to develop a "fully automated risk analysis and intelligence system on fraud and error" continues in stages:

*3.29 Fraud and error systems and reports are not yet complete. **The Department intends to develop a fully automated risk analysis and intelligence system on fraud and error.** But it has not developed this sufficiently to understand and assess fraud and error or to provide staff with effective reporting to enable them to identify potential fraud ... The Department is currently planning to develop the Risk and Intelligence Service **in three stages**: the first stage from April 2018 to October 2018, the second from October 2018 to April 2019, and the third from April 2019 onwards.*

While the Department's plans and timeframes may have changed, its long-held intentions clearly have not.

medConfidential, July 2020

⁵⁹ page 63, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/721224/dwp-annual-report-and-accounts-2017-2018.pdf

⁶⁰ page 61, <https://www.nao.org.uk/wp-content/uploads/2018/06/Rolling-out-Universal-Credit.pdf>